NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

USE OF A BINARY CHOICE MODEL TO DETERMINE MARINE OFFICER ATTRITION

by

Stephen Frederick Hurst

and

Thomas Edward Manion

June 1985

Thesis Advisor: Dan C. Boger

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Use of a Binary Choice Model to Determine Marine Officer Attrition		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis: June, 1985
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Stephen F. Hurst and Thomas E. Manion		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93943-5100		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93943-5100		12. REPORT DATE June, 1985
		13. NUMBER OF PAGES 92
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Marine Officer Attrition, Logit Analysis, Binary Choice

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The ability to predict Marine officer attrition continues to be a significant obstacle to accurate manpower planning at Headquarters Marine Corps. This paper expands on a previous study which used unemployment in a linear regression model to predict attrition. Through the use of a binary choice model, this study incorporates a number of additional variables, principal of which

is a performance index score to measure promotion potential. The logit function was employed to evaluate the inherent causal relationships and to further define the ability of statistical analysis techniques to predict officer attrition.

Use of a Binary Choice Model to Determine Marine Officer Attrition

by

Stephen F. Hurst
Lieutenant Colonel, United States Marine Corps
BBA, Wichita State University, 1969

and

Thomas E. Manion
Captain, United States Marine Corps
BS, Widener University, 1977

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
June 1985

ABSTRACT

The ability to predict Marine officer attrition continues to be a significant obstacle to accurate manpower planning at Headquarters Marine Corps. This paper expands on a previous study which used unemployment in a linear regression model to predict attrition. Through the use of a binary choice model, this study incorporates a number of additional variables, principal of which is a performance index score to measure promotion potential. The logit function was employed to evaluate the inherent causal relationships and to further define the ability of statistical analysis techniques to predict officer attrition.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

A. GENERAL

This thesis is a follow-on to a previous study conducted at the Naval Postgraduate School entitled, "Marine Officer Attrition Model". That study, using economic factors and a linear regression model attempted to predict officer attrition by rank. The ability to correctly predict officer attrition has been in the past and continues to be a significant obstacle to accurate promotion planning and budget preparation at Headquarters, U. S. Marine Corps (HQMC). Previous studies have attempted to improve the unreliable "running average" method through the use of multiple variable linear regression. However, their limited predictive ability in determining an individual's decision to leave the Marine Corps suggested that better results could be achieved through the use of additional variables in a binary choice model.

B. BACKGROUND

1. Historical

The ability to accurately predict officer attrition contributes to the efficient management of the budget, promotion boards, and recruiting requirements. HQMC currently

uses a running average method to forecast attrition. This method uses the average annual attrition for the last six years adjusted by the intuitive assumptions of the officer responsible for determining the officer loss rate. This method is reliable only to a certain degree, with some years proving more accurate than others. The "off years" create significant complications for the efficient management of manpower planning. Because of this problem, HQMC requested a statistical study be undertaken to improve their predicting capabilities. The request was referred to the Naval Postgraduate School, and Major W. J. ESMANN, USMC, conducted an analysis which resulted in his thesis entitled, "Marine Officer Attrition Model".

Major Esmann's initial research seemed to indicate a dramatic improvement in the ability to predict total attrition, but he concluded his thesis with a recommendation that further research be conducted using additional variables in a binary choice model.

The first problem encountered by Major Esmann was to narrow the choice of variables which would be used to predict attrition to a workable number. The initial choice of variables was made through interviews and the intuitive decisions of the officer conducting the analysis. At first, three categories of variables were identified as factors which could determine whether an officer would stay on active duty or leave the Marine Corps. These variables were

military pay, the economy (specifically unemployment), and promotion potential. Of these three variables, promotion potential was immediately eliminated because a method to measure that variable had not yet been developed. However, it is significant to note that throughout the study, those officers interviewed felt promotion potential would be a major factor in determining officer attrition.

Next, military pay was studied using factors which showed its comparability with the private sector. A linear regression model was used to provide an insight to their predictive value. The results of this analysis indicated that comparability between military and civilian pay was not a significant factor in an officer's decision to leave the Marine Corps. These results contradicted a study conducted by the Center for Naval Analysis (CNA) which showed that pay comparability was a factor in an officer's decision to remain on active duty. However, the CNA study included all branches of the Armed Forces while Major Esmann's model considered Marine officers alone. The conclusion was that a Marine officer was a different breed of serviceman, not motivated by pay when making his decision to leave the Marine Corps [Ref. 1].

Once pay comparability and promotion potential were eliminated, only the economy remained as a primary factor. To determine which aspects were significant, the possible economic impact on attrition was broken into three areas:

unemployment, consumer price index (CPI), and the Gross National Product (GNP). Using this economic data, initial tests were again run using a linear regression model with both the CPI and GNP proving insignificant as factors in determining officer attrition.

Finally, unemployment was evaluated and segmented into professional-technical and managerial unemployment. These factors were assigned to groups within the data base, aviation officers being classified as professional-technical and ground officers as managerial. Again, a linear regression model was used to determine the significance of this variable.

The data base used for Major Esmann's study was limited for several reasons. First, the data included information from the years 1976 through 1983 only. This was a result of the Marine Corps not being fully automated prior to this period. Data collected previous to 1976 was incomplete and unreliable due to system complications. Second, attrition resulting from the Viet Nam war made it undesirable to use data from that period.

2. Application and results

As previously mentioned, Major Esmann's model was formulated to assist planners in two ways. First, it was designed to predict future manpower levels, an annual process whereby officer attrition determines accession

requirements and, therefore, recruiting goals. From these figures, the manpower budget is more accurately prepared and the budget call improved. Second, with a more accurate prediction of attrition, realistic eligibility zones can be determined, thus eliminating the mismatch between the number of vacancies and the number of officers selected for promotion.

The results of the model, using unemployment as the key variable, produced significant improvements over the old averaging method. Again, the averaging method took the sum of the last six years attrition, by rank, and divided it by six. The value derived was then modified depending on the intuition of responsible officer. TABLE 1 shows the results of the averaging method, Major Esmann's model, the actual figures for FY83, and a prediction (using his model) for FY84. As depicted in TABLE I, a clear improvement can be seen over the averaging method.

Major Esmann's study concluded that unemployment is a significant factor in predicting officer attrition. However, his concluding recommendation centered around adding a promotion potential variable to the unemployment statistics in order to improve the model's forecasting ability [Ref. 2].

TABLE 1

COMPARISON OF MODEL RESULTS

Rank	Avg Method 1983	Regress Model 1983	Actual 1983	% Error Avg 1983	% Error Reg 1983	Regress Model 1984
LtCol	177	152	127	39.4	19.7	152
Major	186	149	146	27.4	2.1	164
Capt	452	377	392	15.3	-3.8	407
1stLt	525	416	396	32.6	5.1	425

C. PLANNED IMPROVEMENTS

With this initial background, the authors set out to expand the previous study and provide HQMC with an improved method for predicting officer attrition. The problem of establishing a variable that could measure promotion potential, one of the key factors missing from Major Esmann's model, was solved using fitness report data provided by HQMC. An explanation of this process is covered in Chapter II of this thesis.

Next, a number of interviews were conducted with officers at HQMC responsible for predicting officer attrition. These interviews provided insight into the effectiveness of previous attempts to use Major Esmann's model. In essence, HQMC had achieved negative results using the Marine Officer Attrition Model during FY1984 and had reverted to the

averaging method which was proving more accurate. HQMC felt that although 1984 was not a typical year in terms of attrition, the averaging method provided a more accurate prediction than the Esmann model. These discussions gave some initial concerns but the authors believed the use of the performance index would prove to be critical factor in improving attrition prediction.

In the subsequent chapters, the authors will describe the basic data as supplied by HQMC and the transformations required to create the working data base for this study; a description of the statistical functions available for this type of research, their strengths and weaknesses, and the eventual selection of the model that was used; the methodology employed during the data manipulation phase; and an analysis of the results with concluding comments.

II. DATA BASE

A. SOURCE

The data for this thesis were obtained from the Marine Corps Manpower Management System (MMS) and fitness report input. It consisted of 132,903 records containing nineteen data elements on every Marine officers on active duty from 1977 through 1984. A record was created for each officer for each year, so that of the 132,903 records, somewhat less were actually in the data base due to multiple entries for those officers that remained on active duty during the period covered. A detailed example of the data supplied by HQMC is contained in Figure 2.1. The variables included were:

(1) ID CODE. This is an encoded identification of the officer whose data appears on this line.

(2) PMOS. This numeric discriptor indicates an officer's Primary Military Occupational Specialty.

(3) FSTAMOS. This is the First Additional Military Occupational Speciality, or secondary skill the officer is qualified for.

(4) ED MAJ. This a numeric designator describing the officer's civilian education major.

(5) RANK. This alphanumeric variable represents the pay grade equivalent of the officer's rank.

(6) ADBDY, ADBDM, ADBDD. Combined, these variables indicate the Active Duty Base Date. This date is often not the actual date an officer first came on active duty, but has been adjusted for time spent in the reserves or as an enlisted person.

(7) FCOMMY, FCOMMM, FCOMMD. This is the date the officer was first commissioned.

(8) SCORE. This is the officer's Performance Index Score, to be described later in this chapter.

(9) POS SCR. This is the maximum, or Possible Score, that an officer could have attained for his length of service.

(10) DATESEPY, DATESEPM, DATESEPD. This is the date the officer was separated from active service. A "." in these columns means the officer is still on active duty.

(11) CYC YR. This is Manpower Management System date indicating the end of the record year this data was compiled from.

The first change made to the basic data supplied was to eliminate those categories of officers whose presence could bias the results or that would have negligible impact on the outcome. To this end, all second lieutenants were eliminated because they would be promoted prior to reaching a decision point in their career. Those officers with more than four years of enlisted service, warrant officers, and limited duty officers were also eliminated because a preliminary statistical analysis indicated they tended to be career oriented with virtually none leaving the service prior to their twentieth year. The number of officers, by rank, remaining in the data base are displayed in Table II.

In addition to the personal data provided through the MMS, data pertinent to an individual's decision to remain on active duty were obtained from a number of other sources.

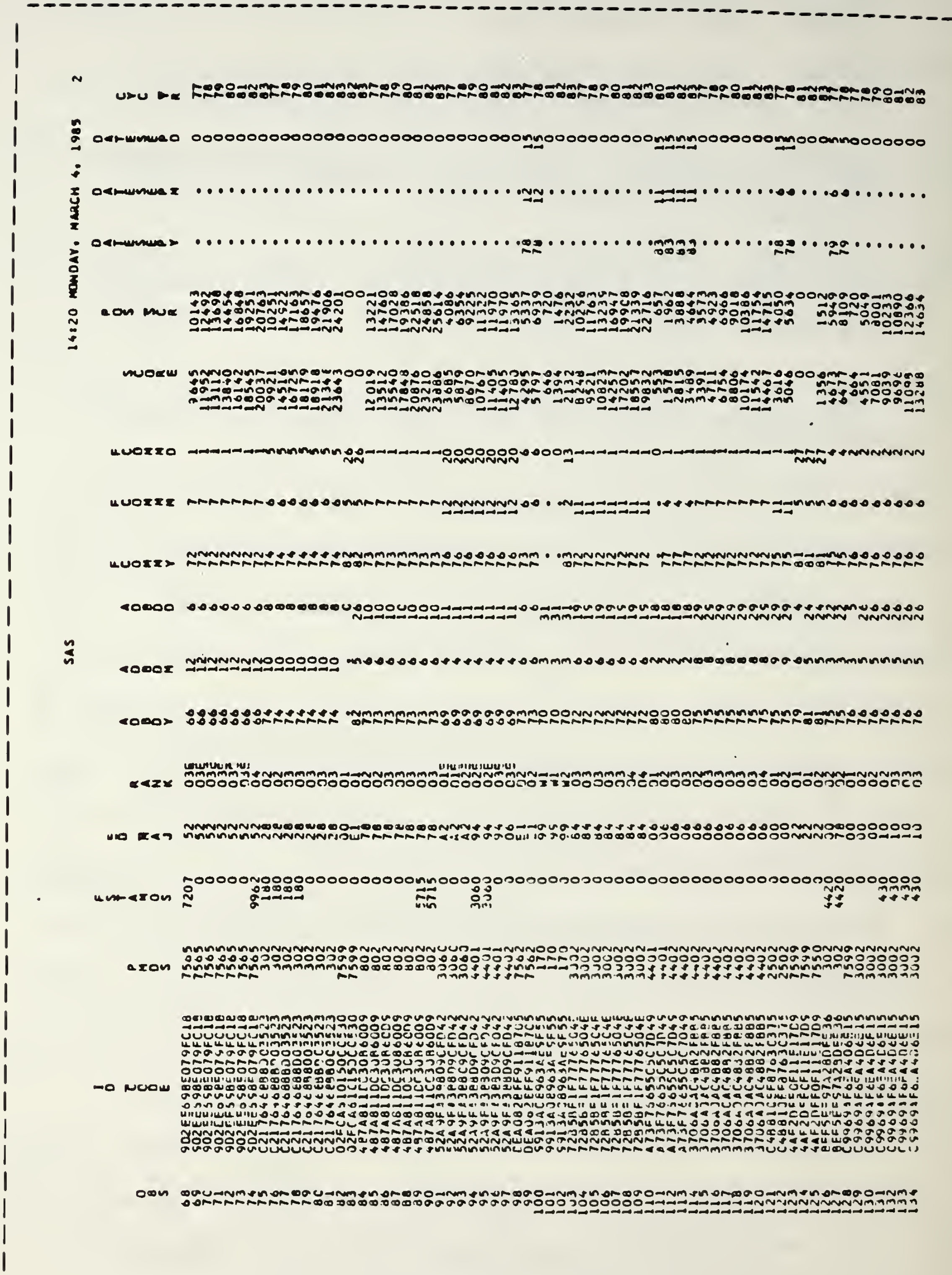


Figure 2.1 Sample of the Basic Data.

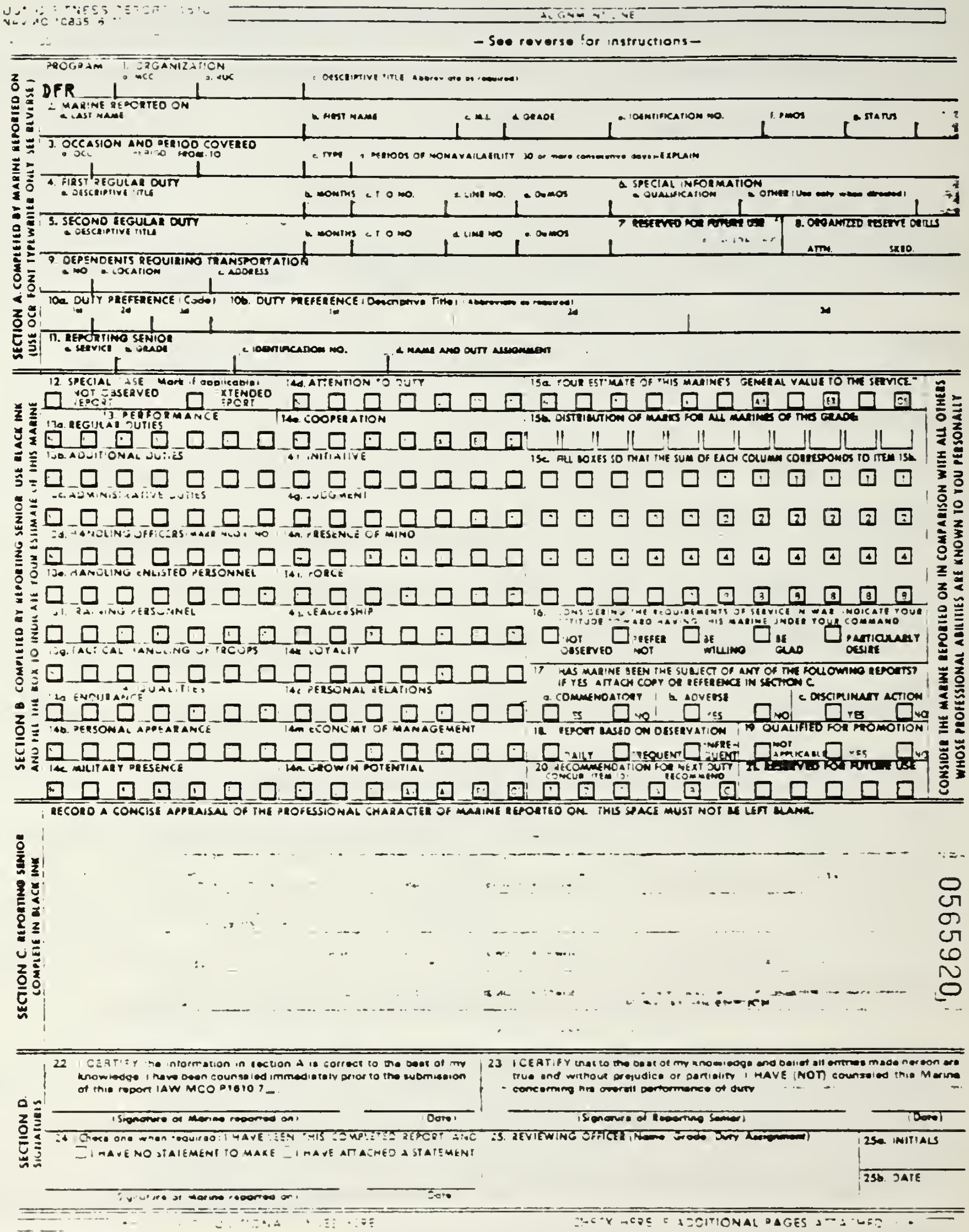
TABLE II

RANK REPRESENTATION IN DATA BASE

Colonel	4,072
Lieutenant Colonel	10,327
Major	20,532
Captain	26,750
First Lieutenant	26,250
TOTAL	87,931

First, from the officer's fitness reports, HQMC provided a performance SCORE which was determined by weighting each of the "Performance" and "Qualities" blocks (see Figure 2.2) of the report and taking this value times a numeric value assigned to the mark received (Outstanding, Excellent, etc.). The maximum score assigned, or POS_SCR, is the score that would have resulted had the officer received a mark of "Outstanding" in each of the observed areas. Both SCORE and POS_SCR are cumulative, with senior officers having a higher value for these variables than a more junior officer.

Second, a relationship between military and civilian pay, expressed as a percentage difference of the former from the latter, was obtained from the government's Employment Cost Index [Ref. 3] and added to the data base for each year of the period covered. The figures used in our analysis are displayed in Table III.



— See reverse for instructions —

SECTION A. COMPLETED BY MARINE REPORTED ON (USE OCR FONT TYPEWRITER ONLY SEE REVERSE)

PROGRAM | 1. ORGANIZATION a. MCC b. RUC | c. DESCRIPTIVE TITLE

DFR

2. MARINE REPORTED ON a. LAST NAME | b. FIRST NAME | c. M.I. | d. GRADE | e. IDENTIFICATION NO. | f. PMOS | g. STATUS

3. OCCASION AND PERIOD COVERED a. OCC | b. PERIOD FROM: TO | c. TYPE | d. PERIODS OF NONAVAILABILITY 30 or more consecutive days—EXPLAIN

4. FIRST REGULAR DUTY a. DESCRIPTIVE TITLE | b. MONTHS | c. T O NO. | d. LINE NO. | e. DuMOS

6. SPECIAL INFORMATION a. QUALIFICATION | b. OTHER

5. SECOND REGULAR DUTY a. DESCRIPTIVE TITLE | b. MONTHS | c. T O NO | d. LINE NO | e. DuMOS

7. RESERVED FOR FUTURE USE | 8. ORGANIZED RESERVE DRILLS ATTN. SKED.

9. DEPENDENTS REQUIRING TRANSPORTATION a. NO | b. LOCATION | c. ADDRESS

10a. DUTY PREFERENCE (Code) 1st 2d 3d | 10b. DUTY PREFERENCE (Descriptive Title) 1st 2d 3d

11. REPORTING SENIOR a. SERVICE | b. GRADE | c. IDENTIFICATION NO. | d. NAME AND DUTY ASSIGNMENT

SECTION B. COMPLETED BY REPORTING SENIOR USE BLACK INK AND FILL THE BOX TO INDICATE YOUR ESTIMATE OF THIS MARINE

12. SPECIAL CASE Mark if applicable: NOT OBSERVED REPORT | EXTENDED REPORT

13. PERFORMANCE: 13a. REGULAR DUTIES | 13b. ADDITIONAL DUTIES | 13c. ADMINISTRATIVE DUTIES | 13d. HANDLING OFFICERS | 13e. HANDLING ENLISTED PERSONNEL | 13f. TRAINING PERSONNEL | 13g. TACTICAL HANDLING OF TROOPS

14. QUALITIES: 14a. ENDURANCE | 14b. PERSONAL APPEARANCE | 14c. MILITARY PRESENCE | 14d. ATTENTION TO DUTY | 14e. COOPERATION | 14f. INITIATIVE | 14g. JUDGMENT | 14h. PRESENCE OF MIND | 14i. FORCE | 14j. LEADERSHIP | 14k. LOYALTY | 14l. PERSONAL RELATIONS | 14m. ECONOMY OF MANAGEMENT | 14n. GROWTH POTENTIAL

15a. YOUR ESTIMATE OF THIS MARINE'S GENERAL VALUE TO THE SERVICE.

15b. DISTRIBUTION OF MARKS FOR ALL MARINES OF THIS GRADE.

15c. FILL BOXES SO THAT THE SUM OF EACH COLUMN CORRESPONDS TO ITEM 15b.

16. CONSIDERING THE REQUIREMENTS OF SERVICE IN WAR, INDICATE YOUR ATTITUDE TOWARD HAVING THIS MARINE UNDER YOUR COMMAND: NOT OBSERVED | PREFER NOT | BE WILLING | BE GLAD | PARTICULARLY DESIRE

17. HAS MARINE BEEN THE SUBJECT OF ANY OF THE FOLLOWING REPORTS? IF YES ATTACH COPY OR REFERENCE IN SECTION C. a. COMMENDATORY YES NO | b. ADVERSE YES NO | c. DISCIPLINARY ACTION YES NO

18. REPORT BASED ON OBSERVATION: DAILY | FREQUENT | INFREQUENT

19. QUALIFIED FOR PROMOTION: NOT APPLICABLE | YES | NO

20. RECOMMENDATION FOR NEXT DUTY CONCUR ITEM 10 RECOMMEND

21. RESERVED FOR FUTURE USE

CONSIDER THE MARINE REPORTED ON IN COMPARISON WITH ALL OTHERS WHOSE PROFESSIONAL ABILITIES ARE KNOWN TO YOU PERSONALLY

SECTION C. REPORTING SENIOR COMPLETE IN BLACK INK

RECORD A CONCISE APPRAISAL OF THE PROFESSIONAL CHARACTER OF MARINE REPORTED ON. THIS SPACE MUST NOT BE LEFT BLANK.

0565920

SECTION D. SIGNATURES

22. I CERTIFY the information in section A is correct to the best of my knowledge. I have been counseled immediately prior to the submission of this report IAW MCO P1610.7_.

(Signature of Marine reported on) (Date)

23. I CERTIFY that to the best of my knowledge and belief all entries made hereon are true and without prejudice or partiality. I HAVE (NOT) counseled this Marine concerning his overall performance of duty.

(Signature of Reporting Senior) (Date)

24. (Check one when required) I HAVE SEEN THIS COMPLETED REPORT AND I HAVE NO STATEMENT TO MAKE | I HAVE ATTACHED A STATEMENT

(Signature of Marine reported on) (Date)

25. REVIEWING OFFICER (Name, Grade, Duty Assignment) | 25a. INITIALS | 25b. DATE

CHECK HERE IF ADDITIONAL PAGES ATTACHED

Figure 2.2 Example of USMC Fitness Report

TABLE III

MILITARY PAY COMPARABILITY

Year	Difference
1977	-4.8
1978	-4.3
1979	-6.5
1980	-7.3
1981	-4.8
1982	-0.4
1983	-4.2

Last, managerial and technical executive unemployment figures were obtained from the Bureau of Labor and Statistics. These figures [Ref. 4], which were lagged by six months to bring them in line with a date closer to the time an officer made the decision to resign, are displayed in TABLE IV.

B. TRANSFORMATIONS

Although some of the data could be used directly, many of the required elements needed to be derived from the basic data as supplied.

An individual's performance INDEX was determined by dividing his SCORE by his POS_SCR. Due to the inflation of fitness reports, the INDEX covered a very narrow range.

TABLE IV

LAGGED UNEMPLOYMENT RATES

Year	Managerial	Tech. Executive
1977	.0295	.0310
1978	.0245	.0280
1979	.0210	.0250
1980	.0225	.0245
1981	.0255	.0265
1982	.0315	.0305
1983	.0355	.0320

The inflation of fitness reports during the period also required the INDEX be normalized to a common mean. This was accomplished using the following formula:

$$INDEXn = (INDEX - INDEXrm) / Std\ Dev$$

where:

INDEXn = Normalized Index

INDEX = given

INDEXrm = Index mean by rank and years of service

Std Dev = Standard Deviation corresponding to the INDEXrm

To determine if an individual's education major played a significant role in his decision to remain on active duty, the Marine Corps Manpower Management System(MMS)

alphanumeric code for education major was converted to a binary (1 and 2) choice that could be used with the LOGIST procedure of SAS. This was accomplished in a somewhat subjective manner based on the marketability of the individual's level and type of degree. As an example, those individuals possessing degrees in the soft skills, such as business administration and education were given an EDSPEC of 1 while those with more technical degrees such as computer science or electrical engineering were given an EDSPEC of 2.

The variable OUT was derived from the DATESEPY element of the basic data. If an individual had a entry for this element, it indicated he had been released from active duty and was assigned an OUT variable value of 1. An individual remaining on active duty received a value of Ø.

YEARS, or the number of years each officer had been on active duty, was simply his active duty base date (ADBDY) minus the data cycle year (CYC_YR), a date used by the MMS to mark the annual update of individual's record.

Because colonels are assigned a 99XX series PMOS, or numerical designator for an officer's primary military occupational specialty, upon promotion, it was necessary to create a new variable entitled MOS that would correlate the actions of these officers with those of lesser grades. For lieutenant colonel and below, MOS equalled PMOS, but for

colonels, MOS equalled their first additional MOS (FSTAMOS), or the MOS they had been assigned prior to promotion.

Finally, a determination was required to accurately present the unemployment statistics, or the UNEMP variable. Like Major Esmann, the authors assigned Naval Aviators and Naval Flight Officers the unemploment figures associated with technical executives while the other officers were assigned the managerial statistics. This was clearly an subjective decision, but one which appeared to most correctly reflect the employability of the population.

An example of the revised data appears in Figure 2.3.

SAS 21:58 FRIDAY, MARCH 15, 1985

OBS	PMOS	FSTAMOS	ED_MAJ	RANK	ADBDY	FCOMMY	DATESEPY	CYC_YR	INDEX	YEARS	EDSPEC	MOS	PAY	UNEMP	OUT	IN
149	3C2	0	76	C3	68	68	100	78	0.906535	10	1	1	-4.5	0.0245	0	1
150	302	0	76	C4	68	68	100	79	0.916114	11	1	1	-6.5	0.0210	0	1
151	3C2	0	76	C4	68	68	100	90	0.921069	12	1	1	-7.3	0.0225	0	1
152	3C2	0	76	C4	68	68	100	81	0.921069	13	1	1	-4.8	0.0255	0	1
153	302	0	76	C4	68	68	100	82	0.922269	14	1	1	-0.4	0.0315	0	1
154	302	0	76	C4	63	68	100	83	0.927319	15	1	1	-4.2	0.0355	0	1
155	302	0	86	C3	67	67	100	77	0.905875	10	1	1	-4.8	0.0295	0	1
156	302	0	86	C3	67	67	100	78	0.917905	11	1	1	-4.5	0.0245	0	1
157	3C2	0	86	C4	67	67	100	79	0.922902	12	1	1	-6.5	0.0210	0	1
158	302	0	N8	C4	67	67	100	80	0.924530	13	1	1	-7.3	0.0225	0	1
159	302	0	N8	C4	67	67	100	81	0.935726	14	1	1	-4.8	0.0255	0	1
160	3C2	0	N8	C4	67	67	100	82	0.940883	15	1	1	-0.4	0.0315	0	1
161	302	0	N8	C5	67	67	100	83	0.944497	16	1	1	-4.2	0.0355	0	1
162	40C2	0	59	C4	66	67	100	77	0.859107	11	1	3	-4.8	0.0295	0	1
163	4002	0	59	C4	66	67	100	78	0.864693	12	1	3	-4.5	0.0245	0	1
164	4002	0	59	C4	66	67	100	79	0.864852	13	1	3	-6.5	0.0210	0	1
165	40C2	0	59	C4	66	67	100	80	0.868996	14	1	3	-7.3	0.0225	0	1
166	4002	0	59	C4	66	67	100	81	0.870810	15	1	3	-4.8	0.0255	0	1
167	40C2	0	59	C5	66	67	100	82	0.865631	16	1	3	-0.4	0.0315	0	1
168	40C2	0	59	C5	66	67	100	83	0.868240	17	1	3	-4.2	0.0355	0	1
169	302	0	86	C3	67	67	100	77	0.877438	10	1	1	-4.8	0.0295	0	1
170	302	0	86	C4	67	67	100	78	0.886859	11	1	1	-4.5	0.0245	0	1
171	302	0	86	C4	67	67	100	79	0.893551	12	1	1	-6.5	0.0210	0	1
172	302	0	86	C4	67	67	100	80	0.895609	13	1	1	-7.3	0.0225	0	1
173	3C2	0	86	C4	67	67	100	81	0.901640	14	1	1	-4.8	0.0255	0	1
174	3C2	0	86	C4	67	67	100	82	0.905560	15	1	1	-0.4	0.0315	0	1
175	302	0	86	C5	67	67	100	83	0.906365	16	1	1	-4.2	0.0355	0	1
176	72C8	0	76	C3	71	71	78	77	0.891475	6	1	3	-4.8	0.0295	1	0
177	4402	0	06	C2	75	72	79	77	0.845574	2	1	3	-4.8	0.0295	0	1
178	44C2	0	06	C3	75	72	79	78	0.794960	3	1	3	-4.5	0.0245	1	0
179	7599	0	64	C2	77	77	83	79	0.817664	2	2	3	-6.5	0.0210	0	1
180	7599	0	64	C2	77	77	83	80	0.944875	3	2	3	-7.3	0.0225	0	1
181	72C8	0	64	C2	77	77	83	81	0.897193	4	2	3	-4.3	0.0255	0	1
182	72C8	0	64	C3	77	77	83	82	0.816919	5	2	3	-0.4	0.0315	1	0
183	25C2	0	52	C2	78	78	100	80	0.897472	2	1	3	-7.3	0.0225	0	1
184	25C2	0	52	C2	78	78	100	81	0.910262	3	1	3	-4.8	0.0255	0	1
185	2502	0	52	C2	78	78	100	82	0.916431	4	1	3	-0.4	0.0315	0	1
186	2502	0	52	C3	78	78	100	83	0.917729	5	1	3	-4.2	0.0355	0	1
187	302	0	78	C2	78	78	82	80	0.915944	2	1	1	-7.3	0.0225	0	1
188	302	0	78	C2	78	78	82	81	0.927988	3	1	1	-4.8	0.0255	1	0
189	3C2	0	00	C2	74	74	100	77	0.877954	3	1	1	-4.8	0.0295	0	1
190	302	0	00	C2	74	74	100	78	0.895008	4	1	1	-4.5	0.0245	0	1
191	302	302	00	C3	74	74	100	79	0.897122	5	1	1	-6.5	0.0210	0	1
192	8C2	302	78	C3	74	74	100	80	0.897919	6	1	1	-7.3	0.0225	0	1
193	302	302	N5	C3	74	74	100	81	0.904327	7	1	1	-4.3	0.0255	0	1
194	3C2	302	N5	C3	74	74	100	82	0.906892	8	1	1	-0.4	0.0315	0	1
195	4C2	302	N5	C3	74	74	100	83	0.920245	9	1	1	-4.2	0.0355	0	1
196	7320	0	59	C2	76	77	100	80	0.874121	4	1	3	-7.3	0.0225	0	1
197	7320	0	59	C2	76	77	100	81	0.892536	5	1	3	-4.8	0.0255	0	1
198	7320	0	59	C2	76	77	100	82	0.907639	6	1	3	-0.4	0.0315	0	1
199	7320	0	59	C3	76	77	100	83	0.915476	7	1	3	-4.2	0.0355	0	1
200	7599	0	26	C2	76	76	81	79	0.829268	3	1	3	-6.5	0.0210	0	1
201	7540	0	26	C2	76	76	81	80	0.848390	4	1	2	-7.3	0.0245	1	0
202	13C2	0	78	C2	76	76	79	79	0.765346	3	1	3	-6.5	0.0210	1	0
203	7562	0	33	C2	77	77	100	80	0.860073	3	2	2	-7.3	0.0245	0	1
204	7562	0	33	C2	77	77	100	81	0.930061	4	2	2	-4.3	0.0265	0	1
205	7562	0	33	C2	77	77	100	82	0.933857	5	2	2	-0.4	0.0305	0	1
206	7562	0	33	C3	77	77	100	83	0.935514	6	2	2	-4.2	0.0320	0	1
207	7599	0		C2	75	75	79	77	0.729571	2	1	3	-4.8	0.0295	0	1
208	13C2	0	29	C2	75	75	79	78	0.743301	3	1	3	-4.5	0.0245	0	1
209	1302	0	29	C2	75	75	79	79	0.743301	4	1	3	-6.5	0.0210	1	0
210	7599	0	22	C2	78	78	100	80	0.842294	2	1	3	-7.3	0.0225	0	1
211	7598	0	22	C2	78	78	100	81	0.892308	3	1	3	-4.3	0.0255	0	1
212	7520	0	22	C2	78	78	100	82	0.835470	4	1	2	-0.4	0.0305	0	1
213	7522	0	22	C2	78	78	100	83	0.887870	5	1	2	-4.2	0.0320	0	1
214	33C2	0	00	C2	76	76	79	78	0.926222	2	1	3	-4.5	0.0245	0	1

Figure 2.3 Sample of the Revised Data Base

III. MODEL SELECTION

A. GENERAL

To assist in the analysis and prediction of Marine officer attrition, it was necessary to select a computer-assisted model that would most appropriately handle the unique aspects of this particular study. Previous studies of the subject had used multivariate regression analysis. Although this procedure was somewhat successful, it was apparent that other models were available that might prove more adept at accurately predicting an officer's decision to remain in the Marine Corps or leave the service. Therefore, following the collection and refinement of the data base, a thorough analysis of the available models was conducted to determine the algorithm that would most accurately predict the binary "stay / leave" decision.

1. Linear Multiple Regression

Linear multiple regression analysis is probably the most commonly used tool of the statistician who wishes to develop a model that will predict a dependent output from a series of known, independent inputs. This form of regression analysis is concerned with modeling the relationship among the independent variables. It quantifies how a response, the dependent variable, is related to a set of explanatory, independent variables. If the true relationship among the variables were known precisely, the

statistician would be able to accurately predict the outcome. However, the true relationship is rarely known and approximations must be developed from empirical evidence.

To accomplish this approximation, the concept of least squares is used to fit the regression equation to the observed data. The least squares method requires that a line be chosen to fit the data so that the sum of the squares of the vertical deviations separating the points from the line will be a minimum. These deviations are represented by the lengths of the vertical line segments that connect the points to the estimated regression line. In the case of multiple regression, the line is replaced by a hyperplane with dimensions equal to the number of independent variables. The resulting estimated multiple regression equation would be:

$$Y = a + b_1 X_1 + b_2 X_2 \ldots + b_n X_n + e$$

where:

Y =	The value of Y calculated from the regression equation.
a =	The value of the constant or "y-intercept".
b_n =	The estimated regression coefficient of a particular independent variable.
X_n =	An independent variable.
e =	Residual error associated with the model

As can be seen, the major goal of linear multiple regression is to predict Y from the regression plane for given levels of X. This is an excellent tool, with some reservations, if the independent variables are continuous and randomly distributed (weight, speed, or payload) and the independent variable will also be continuous, such as cost [Ref. 5]. But, if the investigator wishes to predict the either/or behavior of an individual or group based on the observed past performance of a similar group, then the traditional forms of regression analysis are not able to accurately predict their response or action because the independent variable is qualitative, or at least observed to be qualitative, and not quantitative. Voting, the decision to marry, have children, or go to college are examples of qualitative decisions that are not adequately handled by linear multiple regression. What is needed, therefore, is a set of statistical techniques that can do the work of multiple regression but with qualitative choices as the independent variables. Following will be a discussion of two of the more commonly used procedures to accomplish this task.

2. Discriminant Analysis

Discriminant Analysis deals with a specific class of statistical problem that focuses on the study of groups of populations or data sets. In general it is assumed that the groups being analyzed are discrete and identifiable, that each observation in the group can be described by a set of

measurable variables, and that the variables have a normal distribution within each group. With this in mind, the purpose of discriminant analysis is to test for mean group differences and to describe the overlaps among the groups (the decision to leave the Marine Corps for instance, is a different one than the decision to stay so that the probabilities may add to more than one due to this overlap). Furthermore, discriminant analysis will construct a classification scheme based on the variable set and thereby create a predictive model with which to determine the probable actions of similar groups.

To use discriminant analysis as a predictive tool, a concept similar to regression analysis called the discriminant function is employed. However, unlike the linear regression problem, the dependent variable is dichotomous. This function uses a weighted combination of the independent variables to classify a response into a variable group. Furthermore, using a weighted sum of values derived from the individual predictor variables, a discriminant score is calculated. A cutoff score is then determined that, used in conjunction with the discriminant score, assigns those observations with discriminant scores higher than the cutoff score to one of the criterion groups while those observations with discriminant scores less than the cutoff are assigned to the other [Ref. 6]. The discriminant function is expressed as follows:

$$L = b_1 X_1 + b_2 X_2 \ldots + b_n X_n$$

where:

L = Discriminant score

b_n = Weights associated with the predictors

X_n = Values of the predictor variables

The default cutoff score in most automated systems, such as SAS or SPSS, is .5 with the user being able to modify the breakpoint, if desired, to weight the outcome. This would be done to reduce the number of misclassifications in a particular direction.

In general, discriminant analysis allows the statistician to classify objects into groups based on their values as predictor variables and, much like multiple regression, creates a model that can predict individual or group behavior based on the actions of a similar population base from the past. One of the key factors in accurately predicting behavior, however, is that there be a large difference between the criterion groups. This was certainly not the case with one of the primary variables used in the study of Marine officer prediction, namely the performance INDEX. Due to the inflation of the fitness report marks during the period covered by the data base, it was found that the vast majority of the INDEX values fell into a very narrow range. Thus the difference between an outstanding officer and a less capable one was numerically very small. This led to classification errors and therefore less than

acceptable results when trying to determine what percentage of a particular group would leave the Marine Corps.

3. Logit

Logit is a binary-choice model that assumes that individuals faced with a choice between two alternatives will make a decision based on their respective characteristics. Unlike discriminant analysis, which postulates that the observed values of individual characteristics and attributes of alternatives are drawings from posterior distributions conditioned on actual responses, logit proposes that the actual responses are drawings from multinomial distributions with selection probabilities conditioned on the observed values of individual characteristics and attributes of alternatives. With this in mind, it can be seen that the logit model can be used to calculate the probability of an individual making a particular choice, such as to get out of the Marine Corps, conditioned on an observation of some specific elements of information about that particular officer. To accomplish that process, the function takes the general form:

$$Z_j = \text{Log} \left(\frac{P_j}{1 - P_j} \right) = b_1 X_{1j} + b_2 X_{2j} + \ldots b_n X_{ij} + e$$

where:

Z_j = Probability that a particular decision will be made given a particular variable grouping

P_j = Probability that an individual will make a

particular decision given observation "j" of the variables X

X_{ij} = A variable such that "i" = 1,2,3, etc. representing the 1st, 2nd, and 3rd variables and "j" being the jth observation of a particular variable

b_n = Coefficient associated with each independent variable

To assist in the explanation of this process, the following example of the logit function is provided.

SAMPLE DATA BASE

Name	Sex	Income	Status	P
A	M	H	I	.62
B	M	M	I	.47
C	F	M	I	.69
.	.	.	.	.
.	.	.	.	.
X	M	L	O	.47
Y	M	L	I	.53
Z	M	H	O	.38

P = The percentage of the data base having a particular profile compared to the total number in the independent variable group. For example, males(M) with high(H) income that stayed in(I) divided by males with high income

Applying this percentage to the logit function,

$$Z_j = \text{Log}\left(\frac{P_j}{1 - P_j}\right)$$

the probabilities associated with all (in this example six) of the possible combinations of the independent variables are derived. These are:

Z_1 = Male, high

Z_2 = Male, medium

Z_3 = Male, low

Z_4 = Female, high

Z_5 = Female, medium

Z_6 = Female, low

Once these probabilities have been calculated in conjunction with a particular choice, for instance to get out, the model is run against the population to be tested and the individual probabilities summed to determine the number predicted to leave the Marine Corps. For example, using the following test data,

TEST DATA

Obs	Sex	Income	Z
1	M	H	.34
2	F	M	.69
3	F	H	.52
4	M	L	.23
5	M	L	.23
6	M	M	.38
			2.39

the logit function predicted that 2.39 officers would be leaving the Marine Corps.

Since the logit function makes far fewer assumptions than the discriminant model, such as no multivariate normality for covariates, this form of regression analysis is often preferred to discriminant analysis. In addition, the theory that an officer would leave the Marine Corps

because of pay differences, his educational background, or perceived probability of promotion suggests a causal relationship between the explanatory variables and the probability of his separation. Here again, in side-by-side comparisons, the logit function has proved to be somewhat more accurate in forecasting the direction the dependent variable will take [Ref. 7].

B. MODEL SELECTION

Based on the foregoing analysis, and the availability of the appropriate models at the Naval Postgraduate School, the authors elected to use the logit function to conduct this study. As implemented by the Statistical Analysis System (SAS), the logit function is entitled LOGIST. Supporting the basic program are a number of standard statistical tools to assist the user in the stepwise elimination of insignificant variables. As a guide to the reader in evaluating the SAS LOGIST output, a sample printout is included in Figures 3.1 - 3.4, with the critical points explained below.

1. This matrix depicts the percentile distribution of a particular group based on their individual performance INDEX. This matrix, although not part of the LOGIST procedure, was included to assist in the analysis by displaying the number of officers falling into each of the ten categories. This breakout by INDEX gives the investigator a detailed view of why the model is, or is not, performing as expected. In this example, there is a wide difference between those staying in (the column headed "0") and those leaving (the column headed "1"). Although this difference slowly fades as the INDEX

scores increase, this model should perform very well in predicting officer attrition because of the distinct classification process that has occured.

2. This is the number of observations in the group considered, less those with missing data.

3. Displayed in this area is the basic statistical data on the independent variables.

4. The P-value for testing the joint association of all variables in the model with the dependent variable is displayed at this point. In this example, the P-value of 0 indicates the model is totally significant.

5. These are the P-values of the individual variables deleted from the model as insignificant.

6. Displayed here are the estimated regression coefficients for each variable in the model.

7. This column contains the estimated standard errors of the regression coefficient estimates.

8. This column contains the P-value of the variables retained by the model as being significant.

9. This column depicts the lower confidence limit of the predicted probability.

10. Displayed in this column is the predicted probability of this particular officer leaving the Marine Corps given the values of the independent variables associated with this observation.

11. This column contains the upper confidence limit of the predicted probability.

12. This is the summation of the predicted probabilities for this group, which equals the total number of officers the model predicts will be getting out of the Marine Corps [Ref. 8].

This concludes a description of the model selection procedures. In Chapter IV, the authors will analyze the output resulting from the application of the logit model to each of the groups studied.

LT COLS (GND) WITH 20 YEARS

TABLE OF INTERVAL BY OUT

1

INTERVAL FREQUENCY PERCENT ROW PCT COL PCT	OUT .	0	1	TOTAL
1	11 . . .	54 8.33 83.08 11.66	11 1.70 16.92 5.95	65 10.03
2	12 . . .	55 8.49 84.62 11.88	10 1.54 15.38 5.41	65 10.03
3	21 . . .	55 8.49 84.62 11.88	10 1.54 15.38 5.41	65 10.03
4	22 . . .	69 10.65 82.14 14.90	15 2.31 17.86 8.11	84 12.96
5	11 . . .	33 5.09 70.21 7.13	14 2.16 29.79 7.57	47 7.25
6	13 . . .	44 6.79 67.69 9.50	21 3.24 32.31 11.35	65 10.03
7	19 . . .	48 7.41 73.85 10.37	17 2.62 26.15 9.19	65 10.03
8	21 . . .	39 6.02 60.94 8.42	25 3.86 39.06 13.51	64 9.88
9	16 . . .	36 5.56 55.38 7.78	29 4.48 44.62 15.68	65 10.03
10	11 . . .	30 4.63 47.62 6.48	33 5.09 52.38 17.84	63 9.72
TOTAL	. .	463 71.45	185 28.55	648 100.00

Figure 3.1 Sample SAS Output

```
                         LT COLS (GMD) WITH 20 YEARS
                  STEPWISE LOGISTIC REGRESSION PROCEDURE

                         DEPENDENT VARIABLE: OUT

               548 OBSERVATIONS      2
               403  OUT      = 0
               135  OUT      = 1
               157 OBSERVATIONS DELETED DUE TO MISSING VALUES

   VARIABLE              MEAN         MINIMUM        MAXIMUM         RANGE

   INDXN               0.884749      0.789441       0.983742      0.194301
   PAY         3      -5.77731          -7.3           -4.5           2.8
   UNEMP             0.0244969         0.021         0.0295        0.0085
   EDSPEC              1.16821             1              2             1
   MOS                 1.60494             1              2             1

       -2 LOG LIKELIHOOD FOR MODEL CONTAINING INTERCEPT ONLY=  775.10

STEP  0. THE FOLLOWING VARIABLES ARE ENTERED:

INTERCEP INDXN     PAY       UNEMP      EDSPEC    MOS

         MODEL CHI-SQUARE=   73.46 WITH   5 D.F.   (SCORE STAT.) P=0.0   .
         CONVERGENCE OBTAINED IN  5 ITERATIONS.                  R= 0.297.
         MAX ABSOLUTE DERIVATIVE=0.7961D-06.             -2 LOG L=  696.74.
         MODEL CHI-SQUARE=   78.36 WITH   5 D.F.   (-2 LOG L.R.) P=0.0   . 4

FRACTION OF CONCORDANT PAIRS OF PREDICTED PROBABILITIES AND RESPONSES     :0.707
RANK CORRELATION BETWEEN PREDICTED PROBABILITY AND RESPONSE               :0.430

                         FAST BACKWARD ELIMINATION

   VARIABLE                                  RESIDUAL
   DELETED         CHI-SQUARE       P      CHI-SQUARE      D.F.        P

   EDSPEC             0.03       0.8674  5    0.03          1       0.8674
   PAY                3.05       0.0803       3.03          2       0.2143
   UNEMP              3.84       0.0500       6.92          3       0.0745

STEP  1.

         MODEL CHI-SQUARE=   67.37 WITH   2 D.F.   (SCORE STAT.) P=0.0   .
         CONVERGENCE OBTAINED IN  5 ITERATIONS.                  R= 0.295.
         MAX ABSOLUTE DERIVATIVE=0.7075D-07.             -2 LOG L=  703.81.
         MODEL CHI-SQUARE=   71.29 WITH   2 D.F.   (-2 LOG L.R.) P=0.0   .

FRACTION OF CONCORDANT PAIRS OF PREDICTED PROBABILITIES AND RESPONSES     :0.698
RANK CORRELATION BETWEEN PREDICTED PROBABILITY AND RESPONSE               :0.413

                  TEST FOR ALL VARIABLES NOT IN THE MODEL
              RESIDUAL CHI-SQUARE=    6.99 WITH   3 D.F.  P=0.0723

STEP  2. VARIABLE UNEMP     ADDED WITH CHI-SQUARE=     3.93 P=0.0473 R=-0.050.

         MODEL CHI-SQUARE=   70.39 WITH   3 D.F.   (SCORE STAT.) P=0.0   .
         CONVERGENCE OBTAINED IN  5 ITERATIONS.                  R= 0.299.
         MAX ABSOLUTE DERIVATIVE=0.2317D-06.             -2 LOG L=  699.80.
         MODEL CHI-SQUARE=   75.29 WITH   3 D.F.   (-2 LOG L.R.) P=0.0   .

FRACTION OF CONCORDANT PAIRS OF PREDICTED PROBABILITIES AND RESPONSES     :0.701
RANK CORRELATION BETWEEN PREDICTED PROBABILITY AND RESPONSE               :0.413
```

Figure 3.2 Sample SAS Output

```
LT COLS (END) WITH 20 YEARS
STEPWISE LOGISTIC REGRESSION PROCEDURE

DEPENDENT VARIABLE: OUT

TEST FOR ALL VARIABLES NOT IN THE MODEL
RESIDUAL CHI-SQUARE=     3.09 WITH   2 D.F.   P=0.2135

NO ADDITIONAL VARIABLES MET THE 0.0500 SIGNIFICANCE LEVEL FOR ENTRY.
```

FINAL PARAMETER ESTIMATES

VARIABLE	6 BETA	7 STD. ERROR	CHI-SQUARE	8 P	R
INTERCEPT	-16.21361579	2.59137240	39.15	0.0000	
INDXN	20.52379014	2.91017724	49.74	0.0000	0.248
MOS	-0.97496711	0.19240113	25.68	0.0000	-0.175
UNEMP	-58.57787422	29.62703416	3.91	0.0480	-0.050

Figure 3.3 Sample SAS Output

PARAMETER ESTIMATES AND 95% CONFIDENCE LIMITS

OBS	PMOS	FSTAMOS	EDMAJ	RANK	ADBOY	FCOMMY	DATESEPY	CYC_YR	INDEX	YEARS	EDSPEC	MOS	PAY	UNEMP	OUT	IN	INDXM	INDXN	INTERVAL	9 LOWER_	10 P_	11 UPPER_
1	202	302	10	05	58	58	78	78	0.907608	20	1	1	-4.5	0.0245	1	0	0.6375	0.907608	8	0.4266	0.5013	0.5752
2	802	9644	22	05	61	61	81	81	0.830054	20	1	2	-4.8	0.0255	.	0	-1.5505	0.830054	1	0.0428	0.0678	0.1057
3	302	202	28	05	60	60	100	80	0.898320	20	1	2	-7.3	0.0225	0	1	0.3755	0.898320	7	0.2124	0.2603	0.3146
4	1302	9624	D8	05	57	57	83	77	0.855536	20	2	1	-4.8	0.0295	0	1	-0.8316	0.855536	3	0.1418	0.2046	0.2859
5	7210	7208	E1	05	58	58	78	78	0.899956	20	2	1	-4.5	0.0245	1	0	0.4217	0.899956	7	0.3928	0.4618	0.5322
6	802	0	28	05	60	60	100	80	0.895315	20	1	2	-7.3	0.0225	0	1	0.2907	0.895315	7	0.2021	0.2486	0.3017
7	2502	9624	35	05	60	60	84	80	0.890531	20	2	1	-7.3	0.0225	0	1	0.1558	0.890531	6	0.3731	0.4429	0.5149
8	3060	3002	15	05	57	58	100	77	0.865475	20	2	1	-4.8	0.0295	0	1	-0.5512	0.865475	3	0.1713	0.2398	0.3249
9	802	202	H0	05	59	59	79	79	0.945532	20	1	2	-6.5	0.0210	1	0	1.7075	0.945532	10	0.4028	0.5031	0.6031
10	302	180	40	05	58	58	79	78	0.934216	20	2	2	-4.5	0.0245	0	1	1.3882	0.934216	10	0.3243	0.3953	0.4711
11	7204	7208	55	05	57	56	81	77	0.863149	20	1	1	-4.8	0.0295	0	1	-0.0168	0.863149	3	0.1641	0.2312	0.3154
12	1302	402	52	05	59	59	100	79	0.877763	20	1	1	-6.5	0.0210	0	1	-0.2045	0.877763	4	0.3249	0.4004	0.4810
13	802	5715	61	05	60	60	100	80	0.815469	20	1	2	-7.3	0.0225	0	1	-1.9620	0.815469	1	0.0337	0.0604	0.1004
14	302	9673	52	05	57	58	100	77	0.882106	20	1	2	-4.8	0.0295	0	1	-0.0819	0.882106	5	0.0995	0.1434	0.2023
15	802	9640	04	05	60	60	80	80	0.910737	20	1	2	-7.3	0.0225	1	0	0.7258	0.910737	8	0.2569	0.3122	0.3734
16	302	180	E7	05	57	57	77	77	0.941042	20	2	2	-4.8	0.0295	1	0	1.5803	0.941042	10	0.2646	0.3594	0.4666
17	302	0	28	05	61	60	100	81	0.882575	20	1	2	-4.8	0.0255	.	1	-0.0687	0.882575	5	0.1388	0.1760	0.2206
18	7200	7210	F2	05	61	61	82	81	0.899766	20	1	1	-4.8	0.0255	.	0	0.4163	0.899766	7	0.3763	0.4463	0.5184
19	180	9920	78	05	57	57	92	77	0.838902	20	1	1	-4.8	0.0295	0	1	-1.3009	0.838902	2	0.1006	0.1546	0.2297
20	302	202	73	05	60	60	100	80	0.889416	20	1	2	-7.3	0.0225	0	1	0.1243	0.889416	6	0.1324	0.2256	0.2780
21	3002	0	28	05	60	60	81	80	0.862758	20	1	1	-7.3	0.0225	0	1	-0.0278	0.862758	3	0.2485	0.3101	0.3793
22	302	2502	80	05	57	58	79	77	0.848867	20	1	2	-4.8	0.0295	0	1	-1.0197	0.848867	2	0.0484	0.0780	0.1233
23	302	0	22	05	59	59	79	79	0.897671	20	1	2	-6.5	0.0210	1	0	0.3572	0.897671	7	0.2178	0.2749	0.3404
24	302	202	K6	05	57	57	100	77	0.803339	20	1	2	-4.8	0.0295	0	1	-2.3042	0.803339	1	0.0162	0.0322	0.0628
25	802	0	29	05	60	60	100	80	0.905632	20	1	2	-7.3	0.0225	0	1	0.5813	0.905632	7	0.2363	0.2902	0.3481
26	302	9602	F2	05	61	61	100	81	0.927861	20	1	2	-4.8	0.0255	.	1	1.2090	0.927861	9	0.2870	0.3511	0.4212
27	302	402	28	05	57	60	81	77	0.850942	20	1	2	-4.8	0.0295	0	1	-0.9612	0.850942	2	0.0508	0.0311	0.1272
28	2502	7210	22	05	58	56	81	78	0.870556	20	1	1	-4.5	0.0245	0	1	-0.4078	0.870556	4	0.2623	0.3194	0.3825
29	302	9646	M8	05	60	60	80	80	0.946777	20	1	2	-7.3	0.0225	1	0	1.7426	0.946777	10	0.3952	0.4675	0.5806
30	3002	302	H0	05	57	57	100	77	0.917230	20	1	1	-4.8	0.0295	0	1	0.9090	0.917230	9	0.3705	0.4771	0.5859
31	802	402	29	05	57	58	80	77	0.891004	20	1	2	-4.8	0.0295	0	1	0.1691	0.891004	6	0.1185	0.1673	0.2310
32	202	302	29	05	60	60	82	80	0.871104	20	1	1	-7.3	0.0225	0	1	-0.3923	0.871104	4	0.2847	0.3479	0.4169
33	302	2305	28	05	60	60	82	80	0.913358	20	1	2	-7.3	0.0225	0	1	0.7998	0.913358	8	0.2666	0.3239	0.3869
34	302	430	06	05	57	57	100	77	0.942244	20	1	2	-4.8	0.0295	0	1	1.6147	0.942244	10	0.2688	0.3651	0.4736
35	1803	9640	01	05	58	58	100	78	0.885563	20	1	1	-4.5	0.0245	0	1	0.0156	0.885563	6	0.3283	0.3897	0.4548
36	302	9644	8C	05	61	61	100	81	0.902359	20	1	2	-4.8	0.0255	.	1	0.4895	0.902359	7	0.1986	0.2428	0.2933
37	302	2502	22	05	60	60	100	80	0.878108	20	1	2	-7.3	0.0225	0	1	-0.1947	0.878108	4	0.1476	0.1886	0.2377
38	302	9960	13	05	59	59	80	79	0.881391	20	1	2	-6.5	0.0210	1	0	-0.1021	0.881391	5	0.1639	0.2135	0.2730
39	302	430	10	05	61	61	100	81	0.892700	20	1	2	-4.8	0.0255	.	1	0.2170	0.892700	7	0.1681	0.2082	0.2550
40	302	0	29	05	61	61	100	81	0.923040	20	1	2	-4.8	0.0255	.	1	1.2140	0.923040	9	0.2876	0.3520	0.4222
41	2502	9624	35	05	60	60	100	80	0.863244	20	2	1	-7.3	0.0225	0	1	-0.0141	0.863244	3	0.2506	0.3123	0.3814
42	7210	7204	F2	05	59	59	100	79	0.842608	20	1	1	-6.5	0.0210	0	1	-1.1963	0.842608	2	0.1794	0.2450	0.3252
43	202	9652	01	05	60	60	81	80	0.866447	20	1	1	-7.3	0.0225	1	0	-0.5237	0.866447	3	0.2643	0.3206	0.3956
44	302	0	52	05	59	61	100	79	0.862416	20	1	2	-6.5	0.0210	0	1	-0.6375	0.862416	3	0.1123	0.1553	0.2108
45	802	402	80	05	61	61	100	81	0.852118	20	1	2	-4.8	0.0255	.	1	-0.9280	0.852118	2	0.0721	0.1026	0.1439
46	302	9602	74	05	61	61	82	81	0.937764	20	1	2	-4.8	0.0255	.	0	1.4883	0.937764	10	0.3233	0.3987	0.4792
47	4402	302	06	05	58	58	78	78	0.928365	20	1	1	-4.5	0.0245	1	0	1.2231	0.928365	9	0.5153	0.6058	0.6896
48	1802	9646	M5	05	57	57	78	77	0.923320	20	2	1	-4.8	0.0295	1	0	1.2219	0.923320	9	0.4181	0.5340	0.6463
49	302	9674	57	05	61	61	100	81	0.907031	20	1	2	-4.8	0.0255	.	1	0.0213	0.907031	8	0.2140	0.2608	0.3138
50	1802	302	86	05	60	57	83	80	0.882170	20	1	1	-7.3	0.0225	0	1	-0.0801	0.882170	5	0.3348	0.4010	0.4712

VARIABLE	LABEL	N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	12 SUM	VARIANCE
_P_	PREDICTED PROBABILITY	157	0.27374870	0.13164282	0.06777126	0.65397102	0.01050624	42.97854519	0.01732983

Figure 3.4 Sample SAS Output

IV. ANALYSIS

A. GENERAL

The initial step of the analysis was to perform a preliminary study of the basic data. Through this review, the authors selected and, if necessary, transformed the independent variables which they felt would prove useful in predicting attrition. Along with this procedure, the dependent variables were transformed from the basic data to reflect a binary choice decision. With this accomplished, the basic data were purged to eliminate biased information. Through these procedures, a data base was created from which eight population groups were selected and evaluated using the SAS LOGIST procedure.

B. INDEPENDENT VARIABLES

As indicated in Chapter III, manipulation of the data was required to create a file which would contain the significant factors to conduct the analysis. Once the data base was transformed into a workable file, the independent variables to be used in the study were selected on the basis of intuitive decisions by the authors.

The first variable selected was INDEX. This variable was the primary reason for conducting this follow-on analysis of officer attrition. However, the binary choice model enabled the authors to add additional independent

variables to the model so that their impact could be evaluated. Several variables were selected from the initial study conducted by Major Esmann. He found the unemployment variable, both technical and managerial, to be highly significant in predicting attrition, and the authors felt it could only strengthen the binary choice model. While unemployment was the only strong predictor in the Esmann study, it was believed that one other variable, PAY comparability, might prove significant when combined with INDEX and UNEMPloyment.

Along with performance index, unemployment (both managerial and technical), and pay comparability, two additional variables were identified as possible factors in predicting officer attrition. First, education major was selected because of its ability to determine the "employability" of an individual in the civilian sector. The authors felt that those officers with a technical background would have a better opportunity for gaining outside employment while those with non-technical backgrounds would seek the security of the service. Initial perceptions were that this variable would prove applicable only to the junior officers on their initial service committment. The significance of a technical degree would lose its viability as the officer moved further away from an entry level position in a civilian firm. Second, the authors felt that the PMOS of an officer would be an

indication of his career potential in the Marine Corps. Those officers with a PMOS of Ø3Ø2 (infantry) and Ø8Ø2 (artillery) would seem to be taking the "ideal" career path of a Marine ground officer. Both authors have observed that officers without the above specialties did perceive themselves to be less promotable and therefore more likely to attrit when the opportunity arose.

C. DEPENDENT VARIABLE

The dependent variable for this model was entitled "OUT". This variable positioned a 1 in every record which had a separation date and a Ø in those records for which there was no separation date.

D. SCOPE OF STUDY

1. General

In order to remove data that would bias the outcome and to limit the scope of the study to a reasonable number of population groups, two decisions were required prior to the application of the logit function to the data. First, it was determined that a large portion of the population, for which the decision to attrit was either insignificant or not a factor at all, would have to be eliminated from the data base. Second, once this group was deleted from the data, a decision was required to determine which of the many possible rank by years-of-service population groups could be

analyzed in the time allotted and still provide a accurate projection of officer attrition.

2. Biased Data Elimination

In keeping with the Esmann study, the authors eliminated all ground second lieutenants from the analysis because their initial contracts would require them to remain on active duty until after they would have been promoted to first lieutenant. By the same criterion, aviators were also not evaluated until they had reached the rank of captain.

Additionally, the authors eliminated all limited duty officers and warrant officers. As explained in Chapter II, these officers have normally amassed tenure from prior enlisted service and their committment to remain on active duty is strengthened by their selection for promotion to officer status.

3. Model Population Groups

The initial intention was to run the independent variables against the dependent variable "OUT" by rank. However, further study clearly showed that to run this model against a population categorized by rank only would give a very biased result. A captain with 6 years of service faced with the decision to leave the Marine Corps is influenced by different factors than a captain with 9 years of service. As discussed in Chapter II, the variable YEARS was created from the basic data by subtracting an individual's active duty base date (ADBDY) from the data cycle year (CYC YR).

Through this transformation, the variable YEARS provided the authors with the ability to further separate the RANK populations by the number of years an officer had been on active duty. This procedure greatly expanded the number of population groups which could be evaluated. To reduce the complexity of this problem, a program was written to show the frequency of attrition by year and rank. These data proved useful in evaluating the final results, as well as indicating which years by rank were the critical years for attrition.

The authors, using the results of this analysis, determined that a satisfactory evaluation of Marine officer attrition could be accomplished by studying five years by rank groupings. The significant year is that point in an officer's career when the greatest number of his peers leave the Marine Corps. As shown in Table V, the significant years for each rank were: First lieutenants, year 3; captains, year 5; majors, year 11; lieutenant colonels, year 20; colonels, year 26. Also depicted in Table V are the actual attrition figures for FY 81, the year used to establish the validity of the LOGIST model.

Once this decision was reached, the only remaining factor was to determine whether ground and aviation officers should be separated and tested independently. The decision was, like the Esmann model, to test these groups independently. The authors realized that the aviation and

ground communities were a different breed, separated not only geographically but also by career goals.

Summarizing, Table VI shows the population groups that were evaluated along with the independent variables associated with each group.

TABLE V

SIGNIFICANT YEARS BY RANK

Data Group	Significant Years				
	3	5	11	20	26
1st Lt	183	-	-	-	-
Capt	-	118	-	-	-
Maj	-	-	4	-	-
Lt Col	-	-	-	47	-
Col	-	-	-	-	16

As indicated above, a different variable set was utilized for each group. In the case of UNEMPloyment, the complexity of assigning a technical unemployment rate among the ground community required that all ground officers be given the managerial unemployment rate. On the other hand, aviators with their known technical background were assigned the technical unemployment rate. Similarly, MOS was assigned only within the ground community from first lieutenant to lieutenant colonel, because a distinction could be made in these data groups between the primary combat arms MOS's

TABLE VI

GROUP/VARIABLE CORRELATION

Data Group	Variables Used
1st Lt (Gnd)	INDEX, PAY, UNEMP(Mgr) MOS, EDSPEC
Capt (Gnd)	same as 1st Lt (Gnd)
Capt (Air)	INDEX, PAY, UNEMP(Tech) EDSPEC
Maj (Gnd)	same as Capt (Gnd)
Maj (Air)	same as Capt (Air)
Lt Col (Gnd)	same as Capt (Gnd)
Lt Col (Air)	same as Capt (Air)
Col	INDEX, PAY, UNEMP(Mgr) EDSPEC

(Infantry and Artillery) and the remainder of the population. For colonels, the authors eliminated any MOS distinction because these officers have made a critical transition to an executive level where MOS looses its significance. Because the aviation population groups were tested independent of the ground officers, MOS was not a factor and therefore eliminated.

E. OUTPUT

1. General

Once the data groups were established, the authors, using the LOGIST procedure, analyzed the data groups in two

areas. First, the authors fit the models to the FY77 through FY80 data using the LOGIST function. This was accomplished by the LOGIST procedure, using the stepwise mode, which provided statistical data on those variables remaining in the model as well as those deleted because they failed at the .05 level. The estimated models resulting from this procedure provided the output that was then used to produce a probability, or P-value, for each individual to be tested in the FY81 and FY82 forecast years. The summation of these individual probabilities gave the predicted attrition for the population group being analyzed. Comparing the predicted number with the actual figures provided by HQMC, the authors were able to calculate a prediction rate for each population group. For consistency, this prediction rate was calculated as an absolute value, always using the lower of the two numbers (actual or predicted) as the numerator.

Because the primary purpose of this analysis was to evaluate the predictive capabilities of the logit model of Marine officer attrition, prediction rates are presented first. Then the significance of the variables in the estimated models are discussed. To assist the reader, the authors have included in the discussion that follows for each population group, a table ranking the variables in order of significance by P-value of their associated Chi Square tests. In addition, using a procedure unrelated to

LOGIST, interval tables displaying the performance index by percentile were addressed.

2. First Lieutenants (Ground)

a. Prediction Rates

Prediction rates for this group showed varying results. For FY81 actual attrition for ground officers in this data group was 183. LOGIST predicted a total of 186 officers would attrit during FY81 for a prediction rate of 98 percent. Encouraged by these results the authors decided to evaluate FY82 for a more conclusive study. For FY82 the results of the model were far less accurate. Actual attrition for FY82 was 169 while LOGIST provided a forecast of 140. These figures equated to an 83 percent prediction rate or far below the rate achieved with the FY81 model. Figure 4.1 displays 43 observations from the FY81 data group, along with the predicted probability (column P) of each officer leaving the Marine Corps. Below the observations, in the "SUM" column, is a summation of the predicted probabilities, or the total number LOGIST predicts will leave the Marine Corps in FY81 and FY82

b. Significance of Variables

All variables within this population group proved significant at the .05 level. Performance INDEX showed the lowest P-value indicating its high level of significance to the model while PAY was the least significant with a P-value closest to .05. Table VII

PARAMETER ESTIMATES AND 95% CONFIDENCE LIMITS

PMOS	FSTATUS	EDMAJ	RANK	ADODY	FCOMMY	DATESEPY	CY_Y	INDEX	YEARS	EDSPEC	MOS	PAY	UNEMP	OUT	IN	INDXM	INDXN	INTERVAL	LOWER	P	UPPER
302	180	28	02	74	74	100	77	0.967308	3	1	2	-4.8	0.0295	0	1	1.2254	0.967308	10	0.0076	0.0847	0.1056
2502	0	00	02	75	75	78	78	0.895634	3	1	1	-4.5	0.0245	1	0	0.1968	0.895634	6	0.1766	0.2026	0.2293
302	4420	78	02	75	75	79	78	0.798742	3	1	2	-4.5	0.0245	1	0	-1.1840	0.798742	2	0.3814	0.4262	0.4724
3002	0	00	02	76	76	100	79	0.835014	3	1	1	-6.5	0.0210	0	1	0.0455	0.835014	5	0.2252	0.2471	0.2705
7599	0	64	02	77	77	83	80	0.944375	3	2	1	-7.3	0.0225	0	1	0.8985	0.944375	9	0.0667	0.0848	0.1073
2502	0	52	02	73	73	100	81	0.910262	3	1	1	-4.8	0.0255	.	1	0.4053	0.910262	7	0.1455	0.1636	0.1837
302	0	76	02	73	73	82	81	0.927988	3	1	2	-4.8	0.0255	.	0	0.6579	0.927988	8	0.1444	0.1658	0.1897
802	0	0C	02	74	74	100	77	0.877954	3	1	2	-4.8	0.0295	0	1	-0.0552	0.877954	4	0.1407	0.1747	0.2063
7599	0	26	02	76	76	81	79	0.829268	3	1	1	-6.5	0.0210	0	1	-0.7490	0.829268	2	0.3248	0.3542	0.3847
1302	0	78	02	76	76	79	79	0.765346	3	1	1	-6.5	0.0210	1	0	-1.5600	0.765346	1	0.4537	0.4970	0.5404
1302	0	29	02	75	75	79	78	0.743301	3	1	1	-4.5	0.0245	0	1	-1.9741	0.743301	1	0.4587	0.5035	0.5581
7508	0	22	02	78	78	100	81	0.892308	3	1	1	-4.8	0.0255	.	1	0.1494	0.892308	5	0.1682	0.1875	0.2035
3002	0	00	02	76	76	79	79	0.875701	3	1	1	-6.5	0.0210	1	0	-0.0873	0.875701	4	0.2406	0.2634	0.2876
7531	0	22	02	76	76	100	79	0.953743	3	1	1	-6.5	0.0210	0	1	1.0249	0.953743	9	0.1301	0.1484	0.1689
1302	0	22	02	76	76	100	79	0.935572	3	1	1	-6.5	0.0210	0	1	0.7650	0.935572	9	0.1515	0.1708	0.1918
3002	0	04	02	77	77	81	80	0.914588	3	1	1	-7.3	0.0225	1	0	0.4669	0.914588	7	0.1405	0.1663	0.1832
802	430	51	02	75	75	80	78	0.921906	3	1	2	-4.5	0.0245	0	1	0.5712	0.921906	8	0.1607	0.1929	0.2221
2502	0	PC	02	77	77	81	80	0.920789	3	2	1	-7.3	0.0225	1	0	0.5553	0.920789	7	0.0824	0.1037	0.1297
802	0	26	02	78	78	100	81	0.955128	3	1	2	-4.8	0.0255	.	1	1.0447	0.955128	9	0.1140	0.1341	0.1563
2502	0	E1	02	78	78	100	81	0.931330	3	2	1	-4.8	0.0255	.	1	0.7055	0.931330	8	0.0711	0.0900	0.1134
7597	0	C1	02	77	77	100	80	0.910721	3	1	1	-7.3	0.0225	0	1	0.4118	0.910721	7	0.1512	0.1715	0.1935
7599	0	A9	02	76	76	100	79	0.893240	3	1	1	-6.5	0.0210	0	1	0.1623	0.893240	5	0.2120	0.2333	0.2561
4401	3000	52	02	75	75	82	78	0.921198	3	1	1	-4.5	0.0245	0	1	0.5611	0.921198	7	0.1455	0.1674	0.1921
302	0	63	02	76	76	79	79	0.894539	3	1	2	-6.5	0.0210	1	0	0.1812	0.894539	6	0.2357	0.2646	0.2956
2502	0	00	02	76	76	90	79	0.885207	3	1	1	-6.5	0.0210	1	0	0.0432	0.885207	5	0.2240	0.2458	0.2701
1802	0	52	02	75	75	100	78	0.901699	3	1	1	-4.5	0.0245	0	1	0.2332	0.901699	6	0.1703	0.1939	0.2199
100	0	51	02	76	76	80	79	0.898748	3	1	1	-6.5	0.0210	1	0	0.2412	0.898748	6	0.2034	0.2243	0.2468
7531	0	C1	02	77	77	100	80	0.879706	3	1	1	-7.3	0.0225	0	1	-0.0302	0.879706	5	0.1927	0.2157	0.2408
7530	7210	0C	02	75	75	82	78	0.914043	3	1	1	-4.5	0.0245	0	1	0.4592	0.914048	7	0.1541	0.1753	0.2019
2502	0	N6	02	77	77	100	80	0.983216	3	1	1	-7.3	0.0225	0	1	1.4450	0.983216	10	0.0605	0.0959	0.1139
7597	0	87	02	74	74	60	77	0.803922	3	1	1	-4.8	0.0295	0	1	-1.1102	0.803922	2	0.2255	0.2592	0.2985
7593	0	K6	02	73	73	100	81	0.893656	3	1	1	-4.8	0.0255	.	1	0.1636	0.893656	5	0.1604	0.1836	0.2065
100	0	76	02	78	78	100	81	0.867011	3	1	1	-4.8	0.0255	.	1	-0.2111	0.867011	4	0.2044	0.2256	0.2483
7599	0	36	02	76	76	83	79	0.870085	3	1	1	-6.5	0.0210	0	1	-0.1673	0.870085	4	0.2502	0.2736	0.2983
7596	0	76	02	76	76	100	79	0.931124	3	1	1	-6.5	0.0210	0	1	0.7026	0.931124	8	0.1574	0.1767	0.1978
3002	0	59	02	75	75	100	78	0.908018	3	1	1	-4.5	0.0245	0	1	0.3733	0.908018	6	0.1619	0.1850	0.2106
3000	0	00	02	74	74	73	77	0.944742	3	1	1	-4.8	0.0295	1	0	0.8967	0.944742	9	0.0711	0.0873	0.1068
6002	0	23	02	77	77	81	80	0.746184	3	2	1	-7.3	0.0225	1	0	-1.9330	0.746134	1	0.3620	0.3961	0.4353
1303	430	04	02	75	75	100	78	0.904516	3	1	1	-4.5	0.0245	0	1	0.3234	0.904516	6	0.1665	0.1899	0.2157
402	0	28	02	78	78	81	81	0.729238	3	1	1	-4.8	0.0255	.	0	-2.1745	0.729238	1	0.4608	0.5038	0.5507
7208	0	52	02	74	74	77	77	0.823921	3	1	1	-4.8	0.0295	1	0	-0.8252	0.823921	2	0.1938	0.2254	0.2606
3000	0	35	02	75	75	79	79	0.850271	3	2	1	-6.5	0.0210	1	0	-0.4497	0.850271	3	0.1795	0.2172	0.2604
3002	0	22	02	77	77	80	80	0.929233	3	1	1	-7.3	0.0225	1	0	0.6757	0.929233	3	0.1297	0.1484	0.1694

FIRST LIEUTENANT (GND) FY81 PREDICTION RATE RESULTS

N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM	VARIANCE
1033	0.18049565	0.09351526	0.05454416	0.93355817	0.00290959	186.4522039	0.00874510

FIRST LIEUTENANT (GND) FY82 PREDICTION RATE RESULTS

N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM	VARIANCE
952	0.14728132	0.08136296	0.04555475	0.73744500	0.00263764	140.2122381	0.00662319

Figure 4.1 First Lieutenants (Ground)

shows, by their order of significance, each variable and their corresponding P-value and the Chi Square statistic.

Figure 4.2 shows the final parameter estimates of each variable with its associated P - statistic and regression coefficients of each of the significant variables. These results would seem to indicate that the junior officer is most concerned with his level of success during his initial military contract agreement. Yet another, probably more important factor, is the officer's realization that the Marine Corps is not what he expected. This happens quite early in a young officer's career and his performance and attitude drop off quickly relative to that of his peers. This difference in outlook, and therefore performance, is reflected in the officer's INDEX

TABLE VII

VARIABLE SIGNIFICANCE, FIRST LIEUTENANTS

Variable	P - Value	Chi - Square
INDEX	.	241.65
UNEMP	0.0000	28.49
EDSPEC	0.0001	16.14
MOS	0.0310	4.65
PAY	0.0423	4.12

score and a clear distinction is made between the two groups. Because only those officers with the higher scores

LIEUTENANT (GND) WITH 3 YEARS

STEPWISE LOGISTIC REGRESSION PROCEDURE

DEPENDENT VARIABLE: OUT

```
4108 OBSERVATIONS
3179  OUT      = 0
 929  OUT      = 1
1052 OBSERVATIONS DELETED DUE TO MISSING VALUES
```

VARIABLE	MEAN	MINIMUM	MAXIMUM	RANGE
INDXN	0.881927	0.325617	1	0.674383
PAY	-5.88715	-7.3	-4.5	2.8
UNEMP	0.0238633	0.021	0.0295	0.0085
EDSPEC	1.15287	1	2	1
MOS	1.3313	1	2	1

```
-2 LOG LIKELIHOOD FOR MODEL CONTAINING INTERCEPT ONLY= 4392.07
STEP  0. THE FOLLOWING VARIABLES ARE ENTERED:
INTERCEP INDXN    PAY      UNEMP    EDSPEC    MOS

     MODEL CHI-SQUARE=  335.68 WITH   5 D.F.   (SCORE STAT.) P=0.0    .
     CONVERGENCE OBTAINED IN  5 ITERATIONS.                  R= 0.265.
     MAX ABSOLUTE DERIVATIVE=0.9876D-07.               -2 LOG L= 4074.31.
     MODEL CHI-SQUARE=  317.76 WITH   5 D.F.   (-2 LOG L.R.) P=0.0    .

FRACTION OF CONCORDANT PAIRS OF PREDICTED PROBABILITIES AND RESPONSES    :0.654
RANK CORRELATION BETWEEN PREDICTED PROBABILITY AND RESPONSE              :0.335
```

FINAL PARAMETER ESTIMATES

VARIABLE	BETA	STD. ERROR	CHI-SQUARE	P	R
INTERCEPT	10.05289643	0.86808010	134.11	.	
INDXN	-9.20897329	0.59240968	241.65	.	-0.234
PAY	0.09546295	0.04702005	4.12	0.0423	0.022
UNEMP	-99.43144785	18.62692274	28.49	0.0000	-0.078
EDSPEC	-0.48775572	0.12139434	16.14	0.0001	-0.057
MOS	0.17945947	0.08318732	4.65	0.0310	0.025

CLASSIFICATION TABLE

		PREDICTED NEGATIVE	PREDICTED POSITIVE	TOTAL
TRUE	NEGATIVE	3131	43	3179
	POSITIVE	821	108	929
	TOTAL	3952	156	4108

```
SENSITIVITY: 11.6%  SPECIFICITY: 98.5%  CORRECT: 78.8%
FALSE POSITIVE RATE: 30.8%  FALSE NEGATIVE RATE: 20.8%

FRACTION OF CONCORDANT PAIRS OF PREDICTED PROBABILITIES AND RESPONSES    :0.654
RANK CORRELATION BETWEEN PREDICTED PROBABILITY AND RESPONSE              :0.335
```

Figure 4.2 First Lieutenants (Ground)

continue to make the Marine Corps their career, the distinction between those with the greater promotion potential and those who will fail selection or get out is much less later in their careers. Since the dependent variable is the probability of separating from active service, one would expect a negative relationship between the INDEX score and the dependent variable. This negative coefficient clearly indicates a strong negative correlation between a higher INDEX score and an officer's decision to leave active duty.

Following INDEX is the UNEMPloyment variable. This would seem to indicate that many of these young officers initially made the military their choice of employment because of the high unemployment rates they faced upon graduation from college. Three years later, these same officers, faced with a similar decision, opt for the security of a military career, making the high umemployment rate more profound. Supporting this view is the negative coefficient associated with the UNEMPloyment variable, which indicates that as the unemployment rate increases, the probability that an officer will leave active duty decreases.

The educational background of this data group also proved to be highly significant. Again, since the dependent variable is the probability of separating from the Marine Corps and a higher education level would seem to

indicate increased employability, the expected sign of this coefficient would be positive. However, in this case the negative coefficient assigned to the EDSPEC variable indicates that those officers with the more technical, and therefore more saleable, degrees are remaining in the Marine Corps.

The remaining variables, MOS and PAY, were less significant than the above variables but not enough to be deleted from the model. While MOS and PAY are factors in this model it is understandable why they are less significant. In the case of MOS, the junior officer is at a point in his career that the significance of a combat arms MOS has little effect on his outlook regarding career aspirations in the Marine Corps. If it has, the officer is still at a point in his career to allow him the flexiblity to change his MOS to one of the combat arms. Nevertheless, the positive coefficient associated with this variable indicates, as expected, that those officers with the combat arms MOS's are remaining in the Corps. The least significant of all of the variables was PAY. Interestingly enough, the positive coefficient associated with this variable indicates that as pay becomes less comparable with the private sector, attrition decreases. This apparent anomaly is unexplained. Officers in this data group, for the most part, are experiencing their first full-time job. The military offers these junior officers a comparable

salary to their peers in the civilian sector and therefore they perceive their pay to be adequate.

c. Performance Index Interval Table

Figure 4.3 shows ten equal sub-group intervals within the first lieutenant population group. These sub-groups show actual attrition rates for each index interval during the period FY77 through FY80. For this data group only one interval shows an "out" (1) percentage which is greater than the "in" (0) percentage. This is interval 1, or all those officers with a performance index less than or equal to .796879. At the other end of the scale these percentages change, showing that officers with the higher index have the greatest retention rates. This table clearly indicates that for this data group, the Marine Corps is retaining those officers with the higher performance index while those at the extreme bottom are attriting at a significant rate.

3. Captains (Ground)

a. Prediction Rates

Like the previous data group, the prediction rate for captains in the ground community for FY81 was extremely high. Actual attrition for FY81 was 51 officers while LOGIST predicted 50 officers would attrit for a 98 percent prediction rate. As with the previous data group, the model was also evaluated with FY82 data in order to confirm its validity. The FY82 results were far less

LIEUTENANT (GND) WITH 3 YEARS

TABLE OF INTERVAL BY OUT

INTERVAL (FREQUENCY / PERCENT / ROW PCT / COL PCT)	OUT	0	1	TOTAL
INTERVAL 1	0	182 4.41 47.15 5.71	204 4.95 52.85 21.82	386 9.36
INTERVAL 2	.796879	306 7.42 68.46 9.59	141 3.42 31.54 15.08	447 10.84
INTERVAL 3	.840549	332 8.05 76.15 10.41	104 2.52 23.85 11.12	436 10.57
INTERVAL 4	.862793	342 8.29 78.80 10.72	92 2.23 21.20 9.84	434 10.52
INTERVAL 5	.879596	361 8.75 83.76 11.32	70 1.70 16.24 7.49	431 10.45
INTERVAL 6	.893814	332 8.05 80.00 10.41	83 2.01 20.00 8.88	415 10.06
INTERVAL 7	.908022	340 8.24 83.33 10.66	68 1.65 16.67 7.27	408 9.89
INTERVAL 8	.921425	324 7.85 83.51 10.16	64 1.55 16.49 6.84	388 9.41
INTERVAL 9	.935459	354 8.58 86.76 11.10	54 1.31 13.24 5.78	408 9.89
INTERVAL 10	.955157	317 7.68 85.22 9.94	55 1.33 14.78 5.88	372 9.02
TOTAL	1.0	3190 77.33	935 22.67	4125 100.00

Figure 4.3 First Lieutenants (Ground)

encouraging than those for FY81. Attrition during FY82 was 43 officers while LOGIST forecasted that 32 officers would leave the Corps. This equated to a 74 percent prediction rate. Figure 4.4 shows the results of this analysis.

b. Significance of Variables

With this particular data group, only two variables remain in the model at the .05 significance level. As discussed below, these variables, INDEX and EDSPEC, play a significant role in an officer's decision to attrit from the Marine Corps. Here again, the negative coefficient associated with the INDEX variable indicates that officers with higher performance index scores are much more likely to remain on active duty. The remaining variables used (PAY, UNEMPloyment, and MOS) were deleted from the model because they proved to be insignificant at the .05 level. Table VIII shows the variables used to predict attrition in this group ranked from those proving to be the most significant to those having no impact on the model's ability to predict.

The final parameter estimates along with the variables deleted from the model are shown in Figure 4.5. As in the previous data group, INDEX is again the most significant variable. The output from these two data groups would seem to indicate that an officer's performance level plays a major role in his decision to either leave or stay in the Marine Corps. The other variable in the model, EDSPEC, may remain a factor because these officers, with

PARAMETER ESTIMATES AND 95% CONFIDENCE LIMITS

PMOS	FSTAMUS	ED MAJ	RANK	ADBDY	FCDMMY	DATESEPY	CYC YR	INDEX	YEARS	EDSPEC	MOS	PAY	UNEMP	OUT	IN	INDXM	INOXN	INTERVAL	LOWER	P	UPPER
302	180	28	03	74	74	100	79	0.974480	5	1	2	-6.5	0.0210	0	1	1.8262	0.974480	10	0.0405	0.0557	0.0787
3002	0	84	03	72	72	100	77	0.801088	5	2	1	-4.8	0.0295	0	1	-2.2935	0.801083	1	0.3173	0.4191	0.5294
4402	0	06	03	75	72	100	80	0.976491	5	1	1	-7.3	0.0225	0	1	1.8739	0.976491	10	0.0395	0.0557	0.0778
3002	430	10	03	76	76	100	81	0.890353	5	1	1	-4.8	0.0255	.	1	-0.1791	0.890353	4	0.0900	0.1101	0.1371
802	302	00	03	74	74	100	79	0.897122	5	1	2	-6.5	0.0210	0	1	-0.0096	0.897122	5	0.0925	0.1098	0.1299
7583	0	22	03	76	76	100	81	0.965350	5	1	1	-4.8	0.0255	.	1	1.6096	0.965350	10	0.0451	0.0614	0.0830
302	1960	96	03	72	72	100	77	0.861205	5	1	2	-4.8	0.0295	0	1	-0.3619	0.861205	2	0.1217	0.1470	0.1764
1302	0	22	03	76	76	100	81	0.944045	5	1	1	-4.8	0.0255	.	1	1.1040	0.944045	9	0.0575	0.0738	0.0944
1302	1310	22	03	73	73	83	78	0.841481	5	1	1	-4.5	0.0245	0	1	-1.3299	0.841481	1	0.1377	0.1715	0.2116
4401	3060	52	03	75	75	82	80	0.905649	5	1	1	-7.3	0.0225	0	1	0.1928	0.905649	6	0.0856	0.1023	0.1218
7582	0	00	03	72	72	79	77	0.820409	5	1	1	-4.8	0.0295	0	1	-1.8300	0.820409	1	0.1532	0.2012	0.2567
402	1802	52	03	75	75	100	80	0.921194	5	1	1	-7.3	0.0225	0	1	0.5017	0.921194	8	0.0735	0.0898	0.1092
7581	7210	76	03	75	75	82	80	0.925336	5	1	1	-7.3	0.0225	0	1	0.6600	0.925336	8	0.0704	0.0857	0.1053
302	3962	E1	03	73	73	100	73	0.943275	5	2	2	-4.5	0.0245	0	1	1.2043	0.946275	9	0.1158	0.1550	0.2076
3060	0	10	03	73	73	83	73	0.838821	5	1	1	-4.5	0.0245	0	1	-1.3930	0.838821	1	0.1399	0.1751	0.2169
3002	0	59	03	75	75	100	80	0.905962	5	1	1	-7.3	0.0225	0	1	0.2002	0.905962	6	0.0853	0.1020	0.1215
7587	0	E1	03	72	72	78	77	0.793331	5	2	1	-4.3	0.0295	1	0	-2.4595	0.793331	1	0.3277	0.4355	0.5498
3102	0	29	03	72	72	79	77	0.863439	5	1	1	-4.3	0.0295	0	1	-0.3089	0.863439	3	0.1199	0.1444	0.1729
1503	430	04	03	75	75	100	80	0.921222	5	1	1	-7.3	0.0225	0	1	0.5624	0.921222	8	0.0734	0.0897	0.1092
7587	0	33	03	74	74	81	79	0.909951	5	2	1	-6.5	0.0210	0	1	0.2949	0.909951	6	0.1623	0.2075	0.2617
7583	7204	73	03	76	76	100	81	0.844200	5	1	1	-4.3	0.0255	.	1	-1.2654	0.844200	2	0.1355	0.1680	0.2064
3502	302	29	03	73	73	83	78	0.793315	5	1	1	-4.5	0.0245	0	1	-2.4729	0.793315	1	0.1769	0.2448	0.3254
3002	0	00	03	74	74	100	79	0.887103	5	1	1	-6.5	0.0210	0	1	-0.2472	0.887108	4	0.1007	0.1193	0.1407
130	0	16	03	76	76	100	81	0.949016	5	1	1	-4.3	0.0255	.	1	1.2219	0.949016	9	0.0544	0.0707	0.0915
7599	302	52	03	74	74	100	79	0.963700	5	1	1	-6.5	0.0210	0	1	1.5704	0.963700	10	0.0459	0.0623	0.0839
302	0	61	03	72	72	100	77	0.846255	5	2	2	-4.3	0.0295	0	1	-1.2156	0.846255	2	0.2526	0.3215	0.3993
202	0	78	03	76	76	81	81	0.958333	5	1	1	-4.3	0.0255	.	0	1.4430	0.958333	10	0.0469	0.0652	0.0860
3002	0	78	03	75	75	100	80	0.964567	5	1	1	-7.3	0.0225	0	1	1.5362	0.964567	10	0.0456	0.0619	0.0835
802	430	04	03	76	76	82	81	0.967947	5	1	2	-4.8	0.0255	.	1	1.6712	0.967947	10	0.0457	0.0600	0.0818
302	180	94	03	76	76	100	81	0.930060	5	1	2	-4.8	0.0255	.	1	0.7721	0.930060	8	0.0670	0.0832	0.1030
3002	0	22	03	73	73	100	78	0.870997	5	1	1	-4.5	0.0245	0	1	-0.6295	0.870997	3	0.1138	0.1359	0.1616
7210	7277	30	03	75	75	81	80	0.848323	5	1	1	-7.3	0.0225	1	0	-1.1676	0.848323	2	0.1321	0.1627	0.1980
7537	0	84	03	76	76	100	81	0.898761	5	2	1	-4.3	0.0255	.	1	0.0293	0.898761	5	0.1777	0.2253	0.2811
302	0	52	03	76	76	100	81	0.910422	5	1	2	-4.8	0.0255	.	1	0.3061	0.910422	6	0.0818	0.0983	0.1177
4002	0	22	03	76	76	100	81	0.961268	5	1	1	-4.3	0.0255	.	1	1.5127	0.961263	10	0.0473	0.0636	0.0851
130	802	22	03	74	74	79	79	0.784421	5	1	2	-6.5	0.0210	1	0	-2.6840	0.784421	1	0.1671	0.2604	0.3592
802	0	52	03	73	73	84	78	0.880164	5	1	2	-4.5	0.0245	0	1	-0.2096	0.880164	4	0.1014	0.1202	0.1418
302	0	00	03	75	75	100	80	0.947712	5	1	2	-7.3	0.0225	0	1	1.1910	0.947712	9	0.0552	0.0715	0.0923
2502	0	29	03	75	75	100	80	0.962444	5	1	1	-7.3	0.0225	0	1	1.5406	0.962444	10	0.0466	0.0629	0.0845
302	180	29	03	75	75	100	80	0.914359	5	1	2	-7.3	0.0225	0	1	0.4114	0.914359	7	0.0783	0.0947	0.1141
2502	1502	40	03	72	72	80	77	0.852140	5	2	1	-4.8	0.0295	0	1	-1.0769	0.852140	2	0.2441	0.3097	0.3839
302	0	78	03	76	76	83	81	0.917234	5	1	2	-4.8	0.0255	.	1	0.4089	0.917234	7	0.0764	0.0923	0.1122
302	0	00	03	72	72	79	77	0.924345	5	1	2	-4.3	0.0295	1	0	0.6365	0.924345	8	0.0711	0.0874	0.1070

CAPTAINS (GND) FY81 PREDICTION RATE RESULTS

N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM	VARIANCE
434	0.11420945	0.05742213	0.04901201	0.46903796	0.00275635	49.56690134	0.00329730

CAPTAIN (GND) FY82 PREDICTION RATE RESULTS

N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM	VARIANCE
304	0.10462216	0.05467689	0.04747777	0.38374552	0.00313593	31.80593725	0.00298956

Figure 4.4 Captains (Ground)

```
                                                      CAPTAINS (GND) WITH 5 YEARS
                       STEPWISE LOGISTIC REGRESSION PROCEDURE

                             DEPENDENT VARIABLE: OUT

                 1371 OBSERVATIONS
                 1190  OUT     = 0
                  181  OUT     = 1
                  434 OBSERVATIONS DELETED DUE TO MISSING VALUES

     VARIABLE            MEAN          MINIMUM         MAXIMUM          RANGE

     INOXN           0.903236         0.656302        0.994171        0.33789
     PAY             -6.07644             -7.3            -4.5            2.8
     UNEMP          0.0240051            0.021          0.0295         0.0085
     EDSPEC           1.18831                1               2              1
     MOS              1.40481                1               2              1

         -2 LOG LIKELIHOOD FOR MODEL CONTAINING INTERCEPT ONLY= 1069.95

STEP  0. THE FOLLOWING VARIABLES ARE ENTERED:

INTERCEP INOXN     PAY       UNEMP      EDSPEC    MOS

       MODEL CHI-SQUARE=    58.97 WITH   5 D.F.   (SCORE STAT.) P=0.0    .
       CONVERGENCE OBTAINED IN  5 ITERATIONS.                    R= 0.205.
       MAX ABSOLUTE DERIVATIVE=0.1265D-06.              -2 LOG L= 1014.83.
       MODEL CHI-SQUARE=    55.12 WITH   5 D.F.   (-2 LOG L.R.) P=0.0    .

FRACTION OF CONCORDANT PAIRS OF PREDICTED PROBABILITIES AND RESPONSES    :0.636
RANK CORRELATION BETWEEN PREDICTED PROBABILITY AND RESPONSE              :0.316

                         FAST BACKWARD ELIMINATION

     VARIABLE                                   RESIDUAL
     DELETED        CHI-SQUARE         P       CHI-SQUARE      D.F.         P

     PAY               0.36         0.5458        0.36           1       0.5453
     UNEMP             2.67         0.1020        3.04           2       0.2139
     MOS               3.24         0.0720        6.27           3       0.0990

STEP  1.

       MODEL CHI-SQUARE=    52.95 WITH   2 D.F.   (SCORE STAT.) P=0.0    .
       CONVERGENCE OBTAINED IN  5 ITERATIONS.                    R= 0.204.
       MAX ABSOLUTE DERIVATIVE=0.1903D-07.              -2 LOG L= 1021.30.
       MODEL CHI-SQUARE=    48.66 WITH   2 D.F.   (-2 LOG L.R.) P=0.0    .

FRACTION OF CONCORDANT PAIRS OF PREDICTED PROBABILITIES AND RESPONSES    :0.624
RANK CORRELATION BETWEEN PREDICTED PROBABILITY AND RESPONSE              :0.296

                    TEST FOR ALL VARIABLES NOT IN THE MODEL
                RESIDUAL CHI-SQUARE=     6.32 WITH   3 D.F.   P=0.0970

RESIDUAL CHI-SQUARE IS NOT SIGNIFICANT AT THE 0.0500 LEVEL.
                                                      CAPTAINS (GND) WITH 5 YEARS
                       STEPWISE LOGISTIC REGRESSION PROCEDURE

                             DEPENDENT VARIABLE: OUT

                           FINAL PARAMETER ESTIMATES

    VARIABLE          BETA           STD. ERROR       CHI-SQUARE     P         R

    INTERCEPT      5.38333053        1.63809552         10.80     0.0010
    INOXN         -9.30583703        1.82010268         26.14     0.0000   -0.150
    EDSPEC         0.87254379        0.17918430         23.71     0.0000    0.142
```

Figure 4.5 Captains (Ground)

TABLE VIII

VARIABLE SIGNIFICANCE, CAPTAINS (GND)

Variable	P - Value	Chi - Square
INDEX	0.0000	26.14
EDSPEC	0.0000	23.71
MOS	0.0720	6.27
UNEMP	0.1020	3.04
PAY	0.5458	0.36

five years of service, are very employable in the civilian market based on their educational background and job experience. Supporting this idea is the positive coefficient for the EDSPEC variable. Unlike the first lieutenant population group, this positive relationship indicates that those officers with the more technical educations have a higher probability of leaving the service.

A number of variables were deleted from the model. MOS and PAY may have proved insignificant for the same reasons they were less than totally significant in the previous data group. The deletion of the UNEMPloyment variable may explain the reason EDSPEC becomes significant to the model. A technical education, along with five years of experience as a military manager, may make civilian employment more available for this group than the national unemployment rate indicates.

c. Per ormance Index Interval Table

Th table again shows the highest attrition rate occuring at interval 1, or those officers having a performance index less than or equal to .843051. Indicated is that out of 111 officers in this sub-group, 40 have left the service for a 36 percent attrition rate. The remaining intervals are generally within the same range with interval 2 showing an "in" rate of 87.5 percent with a 12.5 percent "out" rate and interval 10 having an 88 percent "in" rate with an 11.9 percent "out" rate. This table clearly shows that those officers with the lowest performance index, interval 1, attrit at the greatest rate. Figure 4.6 shows the interval table with corresponding attrition rates.

4. Captains (Air)

a. Prediction Rates

Unlike the previous data groups analyzed, LOGIST was unsuccessful in predicting attrition for both FY81 and FY82. Actual attrition for FY81 was 67 officers. The LOGIST sum predicted 29 officers would attrit for a 43 percent prediction rate. For FY82 the actual attrition was 26 officers while LOGIST forecasted 6 officers would attrit for a 23 percent prediction rate. Figure 4.7 depicts the predicted attrition rates provided by LOGIST. The reasons for the model's poor performance are discussed following Table IX.

CAPTAINS (GND) WITH 5 YEARS

TABLE OF INTERVAL BY OUT

INTERVAL	OUT	0	1	TOTAL
FREQUENCY PERCENT ROW PCT COL PCT				
INTERVAL 1	0 .343051	71 5.18 63.96 5.97	40 2.92 36.04 22.10	111 8.10
INTERVAL 2	.343051 .862186	91 6.64 87.50 7.65	13 0.95 12.50 7.18	104 7.59
INTERVAL 3	.862186 .878630	115 8.39 85.19 9.66	20 1.46 14.81 11.05	135 9.85
INTERVAL 4	.878630 .891161	125 9.12 88.65 10.50	16 1.17 11.35 8.84	141 10.28
INTERVAL 5	.891161 .900515	98 7.15 88.29 8.24	13 0.95 11.71 7.18	111 8.10
INTERVAL 6	.900515 .911001	132 9.63 94.29 11.09	8 0.58 5.71 4.42	140 10.21
INTERVAL 7	.911001 .920912	117 8.53 91.41 9.83	11 0.80 8.59 6.08	128 9.34
INTERVAL 8	.920912 .934047	140 10.21 88.05 11.76	19 1.39 11.95 10.50	159 11.60
INTERVAL 9	.934047 .950598	154 11.23 88.00 12.94	21 1.53 12.00 11.60	175 12.76
INTERVAL 10	.950598 1.0	147 10.72 88.02 12.35	20 1.46 11.98 11.05	167 12.18
TOTAL		1190 86.80	181 13.20	1371 100.00

Figure 4.6 Captains (Ground)

PARAMETER ESTIMATES AND 95% CONFIDENCE LIMITS 14:21 THURSDAY, MAY 23, 1985

PS	ESTMAJUS	EDMAJ	RANK	ADBDY	FCUMMY	DATESEPY	CYC YR	INDEX	YEARS	EDSPEC	MOS	PAY	UNEMP	OUT	IN	INDXM	INDXN	INTERVAL	LOWER	P	UPPER
7555	0	29	03	75	75	100	80	0.895646	5	1	3	-7.3	0.0245	0	1	0.0747	0.895646	6	0.2141	0.2511	0.3143
7556	0	E1	03	73	73	79	73	0.922840	5	2	3	-4.5	0.0280	0	1	0.7647	0.922840	8	0.0996	0.1516	0.2244
7501	0	A5	03	75	75	100	30	0.973293	5	2	3	-7.3	0.0245	0	1	2.0449	0.973293	10	0.1032	0.1644	0.2419
7562	0	52	03	76	76	81	31	0.926230	5	1	3	-4.8	0.0265	.	0	0.8507	0.926230	9	0.0566	0.0913	0.1431
7556	7207	22	03	75	75	81	30	0.811300	5	1	3	-7.3	0.0245	1	0	-2.0655	0.811300	1	0.4042	0.5087	0.6125
7562	0	22	03	73	73	100	78	0.857639	5	1	3	-4.5	0.0280	0	1	-0.3397	0.857639	2	0.1531	0.2151	0.2936
7562	0	36	03	76	76	33	31	0.803879	5	1	3	-4.8	0.0265	.	1	-2.2538	0.803879	1	0.2109	0.3227	0.4593
7545	0	76	03	76	76	100	81	0.939544	5	1	3	-4.8	0.0265	.	1	1.1836	0.939544	9	0.0473	0.0779	0.1253
7562	0	22	03	73	73	100	78	0.962619	5	1	3	-4.5	0.0280	0	1	1.7741	0.962619	10	0.0464	0.0671	0.1094
7564	0	22	03	73	73	79	78	0.926095	5	1	3	-4.5	0.0280	0	1	0.8473	0.926095	9	0.0678	0.1027	0.1528
7565	0	22	03	75	75	100	30	0.862391	5	1	3	-7.3	0.0245	0	1	-0.7691	0.862391	3	0.2878	0.3505	0.4190
7501	0	E1	03	72	72	77	77	0.833821	5	2	3	-4.8	0.0310	1	0	-1.4940	0.833821	1	0.4268	0.5322	0.6348
7564	7207	76	03	75	75	100	80	0.861378	5	1	3	-7.3	0.0245	0	1	-0.7448	0.861378	3	0.2901	0.3535	0.4225
7564	0	60	03	75	75	100	30	0.953727	5	1	3	-7.3	0.0245	0	1	1.6753	0.953727	10	0.0985	0.1305	0.1861
7501	7577	50	03	75	75	100	30	0.916634	5	2	3	-7.3	0.0245	0	1	0.6072	0.916634	8	0.2164	0.2864	0.3729
7545	7593	52	03	73	73	100	78	0.897110	5	1	3	-4.5	0.0280	0	1	0.1118	0.897110	6	0.0965	0.1422	0.2009
7562	0	29	03	73	72	100	78	0.893411	5	1	3	-4.5	0.0280	0	1	0.0130	0.893411	5	0.1030	0.1480	0.2082
7563	0	H0	03	76	76	100	81	0.924950	5	1	3	-4.8	0.0265	.	1	0.3182	0.924950	8	0.0576	0.0924	0.1449
7564	0	00	03	76	76	83	31	0.855750	5	1	3	-4.8	0.0265	.	1	-0.9376	0.855750	2	0.1505	0.1974	0.2872
7565	7207	01	03	76	76	100	31	0.830556	5	1	3	-4.8	0.0265	.	1	-1.4247	0.830556	1	0.1576	0.2390	0.3449
7564	7577	52	03	75	75	100	30	0.886560	5	1	3	-7.3	0.0245	0	1	-0.1553	0.886560	5	0.2559	0.2840	0.3401
7565	0	76	03	76	76	100	31	0.937747	5	1	3	-4.8	0.0265	.	1	1.1430	0.937747	9	0.0464	0.0795	0.1290
7565	7956	74	03	75	75	81	30	0.942312	5	1	3	-7.3	0.0245	1	0	1.2568	0.942312	9	0.1233	0.1631	0.2125
7562	0	73	03	75	75	100	30	0.874256	5	1	3	-7.3	0.0245	0	1	-0.4080	0.874256	4	0.2012	0.3170	0.3795
7545	0	22	03	72	72	77	77	0.813441	5	1	3	-4.8	0.0310	1	0	-1.8843	0.813441	1	0.3891	0.4800	0.5723
7562	0	73	03	75	75	100	30	0.881446	5	1	3	-7.3	0.0245	0	1	-0.2856	0.881446	4	0.2452	0.2975	0.3556
7562	0	12	03	75	75	100	30	0.939466	5	1	3	-7.3	0.0245	0	1	1.1856	0.939466	9	0.1231	0.1631	0.2175
7565	0	C4	03	76	76	100	31	0.964709	5	1	3	-4.8	0.0265	.	1	1.8271	0.964709	10	0.0331	0.0578	0.0990
7563	4420	74	03	75	75	81	30	0.870507	5	1	3	-7.3	0.0245	1	0	-0.5632	0.870507	3	0.2696	0.3274	0.3909
7562	0	76	03	75	75	100	30	0.891258	5	1	3	-7.3	0.0245	0	1	-0.0367	0.891258	5	0.2236	0.2720	0.3265
7562	7207	52	03	75	75	31	30	0.911723	5	1	3	-7.3	0.0245	1	0	0.4326	0.911723	7	0.1801	0.2255	0.2738
7564	7207	63	03	76	76	81	31	0.876956	5	1	3	-4.8	0.0265	.	0	-0.3996	0.876956	4	0.1039	0.1500	0.2330
7501	7592	26	03	75	75	82	30	0.950450	5	1	3	-7.3	0.0245	0	1	1.4653	0.950450	10	0.1105	0.1494	0.1990
7562	7207	33	03	75	75	81	30	0.871557	5	2	3	-7.3	0.0245	1	0	-0.5365	0.871557	3	0.3275	0.4165	0.5156
7565	0	36	03	76	76	100	81	0.912270	5	1	3	-4.8	0.0265	.	1	0.4965	0.912270	7	0.0679	0.1058	0.1641
7562	0	51	03	76	76	100	31	0.949655	5	1	3	-4.8	0.0265	.	1	1.4451	0.949655	10	0.0410	0.0691	0.1142
7564	0	22	03	73	73	100	78	0.869676	5	1	3	-4.5	0.0280	0	1	-0.5843	0.869676	3	0.1350	0.1904	0.2616
7562	0	A8	03	75	75	100	30	0.863355	5	1	3	-7.3	0.0245	0	1	-0.7446	0.863355	3	0.2857	0.3478	0.4156
7564	0	10	03	75	76	100	31	0.931711	5	1	3	-4.8	0.0265	.	1	0.9898	0.931711	9	0.0526	0.0854	0.1357
7540	0	10	03	72	72	81	77	0.846110	5	1	3	-4.8	0.0310	0	1	-1.1322	0.846110	2	0.3136	0.3935	0.4738
7501	7207	34	03	75	75	100	80	0.923337	5	2	3	-7.3	0.0245	0	1	0.7773	0.923337	8	0.2015	0.2712	0.3545
7501	7592	35	03	76	76	82	81	0.930124	5	2	3	-4.8	0.0265	.	0	0.9496	0.930124	9	0.0766	0.1250	0.1973
7555	0	63	03	76	76	100	81	0.956638	5	1	3	-4.8	0.0265	.	1	1.6731	0.956638	10	0.0361	0.0521	0.1048

CAPTAINS (AIR) FY81 PREDICTION RATE RESULTS

N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM	VARIANCE
240	0.12237694	0.06245207	0.04419849	0.40057957	0.00403126	29.37046520	0.00390026

CAPTAINS (AIR) FY82 PREDICTION RATE RESULTS

N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM	VARIANCE
202	0.03261768	0.01631540	0.01037438	0.09514230	0.00114761	6.58917455	0.00026613

Figure 4.7 Captains (Air)

b. Significance of Variables

In the case of captains (Air), all variables remain in the model with INDEX again proving the most significant. The variables are shown in Table IX with the P-value and the associated Chi Squares.

TABLE IX

VARIABLE SIGNIFICANCE, CAPTAINS (AIR)

Variable	P - Value	Chi - Square
INDEX	0.0000	26.14
PAY	0.0016	10.01
UNEMP	0.0054	7.74
EDSPEC	0.0315	4.62

Figure 4.8 shows the variables and pertinent statistics provided through the LOGIST procedure. While the LOGIST model proved unsuccessful in predicting attrition for this group, an analysis of the variables may give some insight as to its failure. Like the ground community, the aviation element places a high level of significance on performance. However, PAY becomes more significant within this population group. This is reflected by the negative coefficient associated with the PAY variable which shows a direct connection between higher military pay and retention. This may explain the need of the Marine Corps to pay bonus dollars in order to retain these officers. The UNEMPloyment

```
                                                    CAPTAINS (AIR) WITH 5 YEARS
                    STEPWISE LOGISTIC REGRESSION PROCEDURE

                           DEPENDENT VARIABLE: OUT

                 793 OBSERVATIONS
                 583  OUT     = 0
                 210  OUT     = 1
                 244 OBSERVATIONS DELETED DUE TO MISSING VALUES

     VARIABLE            MEAN          MINIMUM          MAXIMUM           RANGE

     INDXN           0.892373         0.694634         0.993156        0.298522
     PAY             -5.95183             -7.3             -4.5             2.8
     UNEMP          0.0269552           0.0245            0.031          0.0065
     EDSPEC           1.25221                1                2               1

        -2 LOG LIKELIHOOD FOR MODEL CONTAINING INTERCEPT ONLY=  916.76

TEP  0. THE FOLLOWING VARIABLES ARE ENTERED:

NTERCEP INDXN     PAY      UNEMP     EDSPEC

      MODEL CHI-SQUARE=   59.30 WITH   4 D.F.   (SCORE STAT.) P=0.0    .
      CONVERGENCE OBTAINED IN  5 ITERATIONS.                      R= 0.237.
      MAX ABSOLUTE DERIVATIVE=0.3551D-08.                -2 LOG L=  857.12.
      MODEL CHI-SQUARE=   59.64 WITH   4 D.F.   (-2 LOG L.R.) P=0.0    .

RACTION OF CONCORDANT PAIRS OF PREDICTED PROBABILITIES AND RESPONSES     :0.651
ANK CORRELATION BETWEEN PREDICTED PROBABILITY AND RESPONSE              :0.325

                           FINAL PARAMETER ESTIMATES

    VARIABLE          BETA            STD. ERROR      CHI-SQUARE      P         R

    INTERCEPT     1.71863564        3.38581687         0.26    0.6117
    INDXN       -12.75030682        2.06195064        38.24    0.0000    -0.199
    PAY          -0.49904236        0.15774012        10.01    0.0016    -0.093
    UNEMP       188.28116708       67.67093328         7.74    0.0054     0.079
    EDSPEC        0.40481351        0.18824171         4.62    0.0315     0.054

                              CLASSIFICATION TABLE

                                   PREDICTED

                              NEGATIVE   POSITIVE        TOTAL
                             -----------------------
                  NEGATIVE  |   572    |    11    |       583
           TRUE             |          |          |
                  POSITIVE  |   184    |    26    |       210
                             -----------------------  ---------
                  TOTAL     |   756    |    37    |       793

            SENSITIVITY: 12.4%  SPECIFICITY: 98.1%  CORRECT: 75.4%
            FALSE POSITIVE RATE: 29.7%  FALSE NEGATIVE RATE: 24.3%

RACTION OF CONCORDANT PAIRS OF PREDICTED PROBABILITIES AND RESPONSES     :0.651
ANK CORRELATION BETWEEN PREDICTED PROBABILITY AND RESPONSE              :0.325
```

Figure 4.8 Captains (Air)

variable remains significant while the variable closest to .05 is EDSPEC. Although EDSPEC remains in the model, it has the least significance compared to the captain (Ground) model. This may be the result of all of the officers in this group having a technical education. One interesting observation is that in this population group, all variables remain in the model. Similar results were achieved within the first lieutenant model. Both models show officers making a decision on whether to remain in the service after their initial contract obligation.

c. Performance Index Interval Table

Figure 4.9 shows the interval table with the corresponding percentages. This interval table, like the ones that precede it, indicates that attrition is greatest in the first interval. Interval 1, depicting those officers with an INDEX score less than or equal to .840127, shows a 53 percent attrition rate in this population sub-group. While attrition is greatest at interval 1, the next three intervals also show a significant attrition rate when compared with the remaining intervals in the table. The combined attrition rate for these three intervals is 31 percent. Again, this table shows attrition is greatest for those with the lower performance index scores.

5. Majors (Air and Ground)

Although the LOGIST procedure was utilized to analyze these two data groups, the attrition rates were too small to

CAPTAINS (AIR) WITH 5 YEARS

TABLE OF INTERVAL BY OUT

INTERVAL FREQUENCY PERCENT ROW PCT COL PCT		OUT 0	1	TOTAL
INTERVAL 1	0	43 5.40 46.24 7.35	50 6.27 53.76 23.58	93 11.67
INTERVAL 2	.840127	55 6.90 69.62 9.40	24 3.01 30.38 11.32	79 9.91
INTERVAL 3	.859267	56 7.03 70.00 9.57	24 3.01 30.00 11.32	80 10.04
INTERVAL 4	.872385	48 6.02 65.75 8.21	25 3.14 34.25 11.79	73 9.16
INTERVAL 5	.884252	61 7.65 82.43 10.43	13 1.63 17.57 6.13	74 9.28
INTERVAL 6	.893682	57 7.15 85.07 9.74	10 1.25 14.93 4.72	67 8.41
INTERVAL 7	.904269	61 7.65 74.39 10.43	21 2.63 25.61 9.91	82 10.29
INTERVAL 8	.915116	54 6.78 79.41 9.23	14 1.76 20.59 6.60	68 8.53
INTERVAL 9	.925926	76 9.54 81.72 12.99	17 2.13 18.28 8.02	93 11.67
INTERVAL 10	.944003	74 9.28 84.09 12.65	14 1.76 15.91 6.60	88 11.04
TOTAL	1.0	585 73.40	212 26.60	797 100.00

Figure 4.9 Captains (Air)

provide a worthwhile analysis. Actual attrition for FY81 for both the air and ground communities was four officers! During FY82 a total of five officers within these data groups left the Marine Corps. A review of majors mith 11 years of service showed similar attrition rates from FY77 through FY83. With this level of attrition, the authors believed a worthwhile analysis of the data was impossible and discontinued any further study of this group.

6. Lieutenant Colonels (Ground)

a. Prediction Rates

The LOGIST procedure provided poor FY81 and FY82 prediction rates for this data group. For FY81, actual attrition was 34 Marine officers while LOGIST predicted 45 officers would leave the Marine Corps, or a 76 percent prediction rate. The results for FY82 were even less promising. Actual attrition was 15 officers compared to a LOGIST prediction of 44, or a 34 percent prediction rate. Figure 4.10 shows the results of this model for both FY81 and FY82.

b. Significance of Variables

Figure 4.11 shows the final parameter estimates for this model. This data group retained two variables in the construction of the model. As with all of the other data groups studied, INDEX was the most significant. The other variable retained in the model was MOS, continuing the trend (discussed below) that in the ground community at

PARAMETER ESTIMATES AND 95% CONFIDENCE LIMITS

PMOS	FSTAMOS	EDMAJ	RANK	ADBDY	FCOMMY	DATESEPY	CYC YR	INDEX	YEARS	EDSPEC	MOS	PAY	UNEMP	OUT	IN	INDXM	INDXN	INTERVAL	LOWER	P̄	UPPER
202	302	10	05	58	53	78	78	0.907608	20	1	1	-4.5	0.0245	1	0	0.6375	0.907608	8	0.4278	0.5020	0.5762
802	9644	22	05	61	61	81	81	0.830054	20	1	2	-4.8	0.0255	.	0	-1.5505	0.830054	1	0.0470	0.0734	0.1126
302	202	28	05	60	60	100	80	0.898320	20	1	2	-7.3	0.0225	0	1	0.3755	0.898320	7	0.1981	0.2401	0.2878
1302	9624	08	05	57	57	83	77	0.855536	20	2	1	-4.8	0.0295	0	1	-0.8316	0.855536	3	0.2045	0.2596	0.3235
7210	7208	E1	05	58	58	78	78	0.899956	20	2	1	-4.5	0.0245	1	0	0.4217	0.899956	7	0.3944	0.4633	0.5337
802	0	28	05	60	60	100	80	0.895315	20	1	2	-7.3	0.0225	0	1	0.2907	0.895315	7	0.1884	0.2292	0.2759
2502	9624	35	05	60	60	84	80	0.890531	20	2	1	-7.3	0.0225	0	1	0.1558	0.890531	6	0.3527	0.4162	0.4828
3060	9002	15	05	57	58	100	77	0.865475	20	2	1	-4.8	0.0295	0	1	-0.5512	0.865475	3	0.2438	0.3002	0.3634
802	202	HC	05	59	59	79	79	0.945532	20	1	2	-6.5	0.0210	1	0	1.7075	0.945532	10	0.3608	0.4515	0.5392
302	180	40	05	58	58	79	78	0.934216	20	2	2	-4.5	0.0245	0	1	1.3882	0.934216	10	0.3240	0.3956	0.4712
7204	7208	55	05	57	56	81	77	0.863149	20	1	1	-4.8	0.0295	0	1	-0.6168	0.863149	3	0.2345	0.2904	0.3537
1302	402	52	05	59	59	100	79	0.877763	20	1	1	-6.5	0.0210	0	1	-0.2045	0.877763	4	0.2962	0.3550	0.4186
302	5715	61	05	60	60	100	80	0.815469	20	1	2	-7.3	0.0225	0	1	-1.9620	0.815469	1	0.0330	0.0556	0.0922
302	9670	52	05	57	56	100	77	0.882106	20	1	2	-4.8	0.0295	0	1	-0.0819	0.882106	5	0.1482	0.1853	0.2293
802	9640	04	05	60	60	80	80	0.910737	20	1	2	-7.3	0.0225	1	0	0.7253	0.910737	8	0.2402	0.2890	0.3433
302	180	E7	05	57	57	77	77	0.941042	20	2	2	-4.8	0.0295	1	0	1.5803	0.941042	10	0.3499	0.4291	0.5120
302	0	28	05	61	60	100	81	0.882575	20	1	2	-4.8	0.0255	.	1	-0.0637	0.882575	5	0.1495	0.1868	0.2308
7208	7210	F2	05	61	61	82	81	0.899766	20	1	1	-4.8	0.0255	.	0	0.4103	0.899766	7	0.3935	0.4523	0.5326
130	9920	78	05	57	57	82	77	0.838902	20	1	1	-4.8	0.0295	0	1	-1.3009	0.838902	2	0.1477	0.2002	0.2655
302	202	73	05	60	60	100	80	0.889416	20	1	2	-7.3	0.0225	0	1	0.1243	0.889416	6	0.1698	0.2087	0.2539
3002	0	28	05	60	60	81	80	0.862758	20	1	1	-7.3	0.0225	0	1	-0.6278	0.862758	3	0.2328	0.2888	0.3521
302	2502	80	05	57	58	79	77	0.848867	20	1	2	-4.8	0.0295	0	1	-1.0197	0.848867	2	0.0731	0.1039	0.1456
302	0	22	05	59	59	79	79	0.897671	20	1	2	-6.5	0.0210	1	0	0.3572	0.897671	7	0.1900	0.2377	0.2852
302	202	K6	05	57	57	100	77	0.803339	20	1	2	-4.8	0.0295	0	1	-2.3042	0.803339	1	0.0245	0.0440	0.0779
802	0	29	05	60	60	100	80	0.905632	20	1	2	-7.3	0.0225	0	1	0.5818	0.905632	7	0.2226	0.2682	0.3193
302	9602	F2	05	61	61	100	81	0.927861	20	1	2	-4.8	0.0255	.	1	1.2090	0.927861	9	0.3013	0.3652	0.4343
302	402	28	05	57	58	81	77	0.850942	20	1	2	-4.8	0.0295	0	1	-0.9612	0.850942	2	0.0767	0.1079	0.1497
2502	7210	22	05	58	56	81	78	0.870556	20	1	1	-4.5	0.0245	0	1	-0.4078	0.870556	4	0.2651	0.3223	0.3854
302	9640	M8	05	60	60	80	80	0.946777	20	1	2	-7.3	0.0225	1	0	1.7426	0.946777	10	0.3715	0.4578	0.5467
3002	302	H0	05	57	57	100	77	0.917230	20	1	1	-4.8	0.0295	0	1	0.9090	0.917230	9	0.4090	0.5206	0.6296
802	402	29	05	57	58	80	77	0.891004	20	1	2	-4.8	0.0295	0	1	0.1691	0.891004	6	0.1747	0.2141	0.2596
202	302	29	05	60	60	82	80	0.871104	20	1	1	-7.3	0.0225	0	1	-0.3923	0.871104	4	0.2674	0.3247	0.3878
302	2305	28	05	60	60	82	80	0.913358	20	1	2	-7.3	0.0225	0	1	0.7998	0.913358	8	0.2493	0.3001	0.3562
302	430	06	05	57	57	100	77	0.942244	20	1	2	-4.8	0.0295	0	1	1.0147	0.942244	10	0.3544	0.4351	0.5193
1803	9640	D1	05	58	58	100	78	0.885563	20	1	1	-4.5	0.0245	0	1	0.0156	0.885563	6	0.3306	0.3920	0.4570
302	9648	80	05	61	61	100	81	0.902359	20	1	2	-4.8	0.0255	.	1	0.4895	0.902359	7	0.2115	0.2554	0.3048
302	2502	22	05	60	60	100	80	0.873108	20	1	2	-7.3	0.0225	0	1	-0.1947	0.878108	4	0.1371	0.1734	0.2169
302	9560	13	05	59	59	80	79	0.881591	20	1	2	-6.5	0.0210	1	0	-0.1021	0.881391	5	0.1461	0.1831	0.2270
302	430	10	05	61	61	100	81	0.892700	20	1	2	-4.8	0.0255	.	1	0.2170	0.892700	7	0.1800	0.2200	0.2659
302	0	29	05	61	61	100	81	0.928040	20	1	2	-4.8	0.0255	.	1	1.2140	0.928040	9	0.3019	0.3600	0.4353
2502	9624	35	05	60	60	100	80	0.863244	20	2	1	-7.3	0.0225	0	1	-0.6141	0.863244	3	0.2347	0.2908	0.3541
7210	7204	F2	05	59	59	100	79	0.842608	20	1	1	-6.5	0.0210	0	1	-1.1963	0.842608	2	0.1593	0.2125	0.2776
202	9652	D1	05	60	60	81	80	0.866447	20	1	1	-7.3	0.0225	1	0	-0.5237	0.866447	3	0.2478	0.3044	0.3675

LIEUTENANT COLONEL (GND) FY81 PREDICTION RATE RESULTS

N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM	VARIANCE
157	0.28632836	0.13375364	0.07335478	0.66621810	0.01067470	44.95355219	0.01789004

LIEUTENANT COLONEL (GND) FY82 PREDICTION RATE RESULTS

N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM	VARIANCE
151	0.29406664	0.13471667	0.05781548	0.66938615	0.01096309	44.40406232	0.01814858

Figure 4.10 Lieutenant Colonels (Ground)

```
                                                  LT COLS (GND) WITH 20 YEARS
                  STEPWISE LOGISTIC REGRESSION PROCEDURE

                          DEPENDENT VARIABLE: OUT

              648 OBSERVATIONS
              463  OUT      = 0
              185  OUT      = 1
              157 OBSERVATIONS DELETED DUE TO MISSING VALUES

   VARIABLE              MEAN          MINIMUM         MAXIMUM          RANGE

   INOXN             0.884749         0.789441        0.983742       0.194301
   PAY               -5.77731             -7.3            -4.5            2.8
   UNEMP           0.02443069            0.021          0.0295         0.0085
   EDSPEC             1.16821                1               2              1
   MOS                1.60494                1               2              1

        -2 LOG LIKELIHOOD FOR MODEL CONTAINING INTERCEPT ONLY=  775.10

STEP  0. THE FOLLOWING VARIABLES ARE ENTERED:

INTERCEP INOXN     PAY       UNEMP      EOSPEC     MOS

        MODEL CHI-SQUARE=    73.46 WITH    5 D.F.    (SCORE STAT.) P=0.0    .
        CONVERGENCE OBTAINED IN  5 ITERATIONS.                     R= 0.297.
        MAX ABSOLUTE DERIVATIVE=0.7901D-06.               -2 LOG L=  696.74.
        MODEL CHI-SQUARE=    78.36 WITH    5 D.F.    (-2 LOG L.R.) P=0.0    .

FRACTION OF CONCORDANT PAIRS OF PREDICTED PROBABILITIES AND RESPONSES    :0.707
RANK CORRELATION BETWEEN PREDICTED PROBABILITY AND RESPONSE              :0.430

                      FAST BACKWARD ELIMINATION

   VARIABLE                                 RESIDUAL
   DELETED       CHI-SQUARE          P      CHI-SQUARE      D.F.           P

   EDSPEC           0.03        0.8674         0.03          1        0.8674
   PAY              3.05        0.0803         3.08          2        0.2148
   UNEMP            3.84        0.0500         6.92          3        0.0746

STEP  1.

        MODEL CHI-SQUARE=    67.37 WITH    2 D.F.    (SCORE STAT.) P=0.0    .
        CONVERGENCE OBTAINED IN  5 ITERATIONS.                     R= 0.295.
        MAX ABSOLUTE DERIVATIVE=0.7075D-07.               -2 LOG L=  703.81.
        MODEL CHI-SQUARE=    71.29 WITH    2 D.F.    (-2 LOG L.R.) P=0.0    .

FRACTION OF CONCORDANT PAIRS OF PREDICTED PROBABILITIES AND RESPONSES    :0.698
RANK CORRELATION BETWEEN PREDICTED PROBABILITY AND RESPONSE              :0.413

                 TEST FOR ALL VARIABLES NOT IN THE MODEL
            RESIDUAL CHI-SQUARE=     6.99 WITH    3 D.F.   P=0.0723

RESIDUAL CHI-SQUARE IS NOT SIGNIFICANT AT THE 0.0500 LEVEL.
                                                  LT COLS (GND) WITH 20 YEARS

                          DEPENDENT VARIABLE: OUT

                       FINAL PARAMETER ESTIMATES

  VARIABLE            BETA            STD. ERROR      CHI-SQUARE      P         R

  INTERCEPT   -17.42502086         2.52255510        47.72   0.0000
  INOXN        20.27835291         2.89204992        49.16   0.0000    0.247
  MOS          -0.97168114         0.19163512        25.71   0.0000   -0.175
```

Figure 4.11 Lieutenant Colonels (Ground)

least, having a combat arms MOS is perceived to be important to an officer's future in the Marine Corps. Variables deleted from the model were UNEMP, PAY and EDSPEC. Table X shows the variables for this data group.

While INDEX remained the most significant variable throughout this study, the positive coefficient assigned to INDEX shows a distinct break from the preceding

TABLE X

VARIABLE SIGNIFICANCE, LIEUTENANT COLONELS (GND)

Variable	P - Value	Chi - Square
INDEX	0.0000	47.72
MOS	0.0000	25.71
UNEMP	0.0500	3.84
PAY	0.0808	3.05
EDSPEC	0.8678	0.03

population groups. This indicates, for the first time in this study, that those officers with the higher performance index scores have a greater probability of attrition. This output supports the data displayed in the Performance Index Interval Table discussed below. Additionally this data group placed a high level of significance on MOS. This would seem to indicate that after reaching the rank of lieutenant colonel, officers begin to regard their job assignments as a means of enhancing their future chances for

promotion. Marine officers in this data group with combat arms MOS's have had more command assignments, making their future promotion potential greater than those who have never served in this type of job due to their restrictive MOS's.

Addressing the variables deleted, EDSPEC seems to become less significant as the officer moves further away from his undergraduate years. This supports the authors' original impressions about this variable, believing it would prove significant only for the younger officers. PAY was also deleted. This is because these officers have spent twenty years in the military, and any real concerns with pay would have forced them to attrit at an earlier stage in their careers. UNEMPloyment is just insignificant at a P-value of .05, which may indicate that the civilian positions this group are seeking are not related to the national unemployment rate.

c. Performance Index Interval Table

Figure 4.12 shows the interval table with the corresponding attrition percentages. With this data group, the performance INDEX does a complete reversal from the previous data groups. In the previous populations, attrition was highest among the intervals representing the lowest index scores. A percentile breakdown of the performance INDEX scores of this data group reveals that those officers with the highest scores attrit at the greatest rate. Interval 1, or those officers with a score

LT COLS (GND) WITH 20 YEARS

TABLE OF INTERVAL BY OUT

INTERVAL FREQUENCY PERCENT ROW PCT COL PCT	OUT	0	1	TOTAL
	0			
INTERVAL 1		54 8.33 83.08 11.66	11 1.70 16.92 5.95	65 10.03
	.838258			
INTERVAL 2		55 8.49 84.62 11.88	10 1.54 15.38 5.41	65 10.03
	.855389			
INTERVAL 3		55 8.49 84.62 11.88	10 1.54 15.38 5.41	65 10.03
	.866850			
INTERVAL 4		69 10.65 82.14 14.90	15 2.31 17.86 8.11	84 12.96
	.873684			
INTERVAL 5		33 5.09 70.21 7.13	14 2.16 29.79 7.57	47 7.25
	.884532			
INTERVAL 6		44 6.79 67.69 9.50	21 3.24 32.31 11.35	65 10.03
	.892454			
INTERVAL 7		48 7.41 73.85 10.37	17 2.62 26.15 9.19	65 10.03
	.905655			
INTERVAL 8		39 6.02 60.94 8.42	25 3.86 39.06 13.51	64 9.88
	.916338			
INTERVAL 9		36 5.56 55.38 7.78	29 4.48 44.62 15.68	65 10.03
	.929695			
INTERVAL 10		30 4.63 47.62 6.48	33 5.09 52.38 17.84	63 9.72
	1.0			
TOTAL		463 71.45	185 28.55	648 100.00

Figure 4.12 Lieutenant Colonels (Ground)

less than or equal to .838258, have an "in" percentage of 83 percent and an "out" percentage of 16.9. On the other hand, interval 10 has an "out" rate of 52 percent and an "in" rate of 47.6 percent. A review of this table clearly shows that those officers in the higher interval subgroupings attrited at the greatest rate. This would indicate the Marine Corps is losing those officers with the higher performance INDEX scores from this data group.

7. Lieutenant Colonel (Air)

a. Prediction Rates

The application of LOGIST to this population group produced mixed results for FY81 and FY82. The prediction rate for FY81 was 93 percent, with LOGIST forecasting 14 officers against an actual attrition of 13 officers. For FY82, LOGIST predicted 13 officers would leave the Marine Corps when actually 16 officers resigned their commissions. This resulted in a predictability rate of only 81 percent for FY82. Figure 4.13 displays the LOGIST output for this data group.

b. Significance of Variables

INDEX was the only variable that remained in this model. The input variables for this model were similar to those used to analyze the lieutenant colonel (Ground) population group with the exception of MOS. MOS was not used with this group because, like the other groups involving pilots, they had been arranged by MOS initially.

PARAMETER ESTIMATES AND 95% CONFIDENCE LIMITS

PMOS	FSTAMCS	ED MAJ	RANK	ADBDY	FCOMMY	DATESEPY	CYC_YR	INDEX	YEARS	EDSPEC	MOS	PAY	UNEMP	OUT	I_N	INDXM	INDXN	INTERVAL	LOWER_	P_	UPPER_
7522	7596	N5	05	59	59	80	79	0.932495	20	1	3	-6.5	0.0250	1	0	1.2714	0.932495	9	0.2653	0.3623	0.4713
7501	7560	H0	05	57	57	78	77	0.944601	20	1	3	-4.3	0.0310	0	1	1.0415	0.944601	10	0.2970	0.4239	0.5607
7522	7530	52	05	57	57	79	77	0.906262	20	1	3	-4.8	0.0310	0	1	0.4694	0.906262	7	0.1902	0.2448	0.3092
7562	7564	10	05	61	61	81	81	0.873386	20	1	3	-4.8	0.0265	.	0	-0.5358	0.873386	4	0.0923	0.1383	0.2023
7522	7555	15	05	61	62	82	81	0.870668	20	2	3	-4.8	0.0265	.	1	-0.6189	0.870668	3	0.0858	0.1316	0.1964
7562	7565	78	05	60	61	100	80	0.886127	20	1	3	-7.3	0.0245	0	1	-0.1462	0.886127	5	0.1270	0.1741	0.2341
7563	7207	22	05	61	61	100	81	0.900739	20	1	3	-4.8	0.0265	.	1	0.3020	0.900739	7	0.1729	0.2239	0.2848
7565	7501	64	05	58	58	80	78	0.853847	20	2	3	-4.5	0.0230	0	1	-1.1332	0.853847	2	0.0535	0.0955	0.1650
7522	7501	N5	05	60	60	81	80	0.853344	20	1	3	-7.3	0.0245	1	0	-0.9956	0.853344	2	0.0609	0.1043	0.1727
7522	7521	E1	05	59	59	100	79	0.852613	20	2	3	-6.5	0.0250	0	1	-1.1709	0.852613	2	0.0516	0.0934	0.1630
7522	7501	D1	05	57	57	100	77	0.912951	20	1	3	-4.8	0.0310	0	1	0.6739	0.912951	9	0.2107	0.2722	0.3440
7511	7595	M8	05	60	60	100	80	0.906301	20	1	3	-7.3	0.0245	0	1	0.4706	0.906301	7	0.1903	0.2450	0.31[illegible]
7562	7565	16	05	61	61	100	81	0.875648	20	1	3	-4.8	0.0265	.	1	-0.4666	0.875648	4	0.0979	0.1442	0.2[illegible]
7557	7562	48	05	60	60	100	80	0.874435	20	1	3	-7.3	0.0245	0	1	-0.5037	0.874435	4	0.0947	0.1410	0.20[illegible]
7501	7620	31	05	61	62	82	81	0.9[illegible]	20	2	3	-4.8	0.0265	.	0	0.4776	0.906531	7	0.1910	0.2459	0.3105
7564	7562	E5	05	59	59	82	79	0.8[illegible]	20	2	3	-6.5	0.0250	0	1	-1.4299	0.844141	1	0.0402	0.0791	0.1498
7562	7564	15	05	58	59	80	78	0.882433	20	2	3	-4.5	0.0230	0	1	-0.2592	0.882433	5	0.1162	0.1631	0.2240
7565	7563	48	05	60	61	100	80	0.908478	20	1	3	-7.3	0.0245	0	1	0.5371	0.908478	7	0.1971	0.2537	0.3202
7545	5903	35	05	61	61	100	81	0.880181	20	2	3	-4.8	0.0265	.	1	-0.3280	0.880181	4	0.1099	0.1566	0.2182
7564	0	A0	05	61	63	100	81	0.901240	20	1	3	-4.8	0.0265	.	1	0.3158	0.901240	7	0.1743	0.2255	0.2866
7575	7560	52	05	53	53	100	78	0.918089	20	1	3	-4.5	0.0280	0	1	0.8309	0.918089	8	0.2253	0.2945	0.3741
7562	7650	D9	05	60	60	100	80	0.907783	20	2	3	-7.3	0.0245	0	1	0.5159	0.907783	7	0.1949	0.2509	0.3167
7562	7557	G0	05	58	59	100	78	0.891072	20	1	3	-4.5	0.0230	0	1	0.0049	0.891072	6	0.1420	0.1898	0.2491
7564	7522	G2	05	60	60	80	80	0.853818	20	1	3	-7.3	0.0245	1	0	-1.1340	0.853818	2	0.0535	0.0956	0.1650
7563	7565	04	05	61	61	81	81	0.891043	20	1	3	-4.8	0.0265	.	0	0.0042	0.891043	6	0.1417	0.1898	0.2490
7562	7554	10	05	58	59	79	78	0.905273	20	1	3	-4.5	0.0280	0	1	0.4391	0.905273	7	0.1871	0.2410	0.3046
7562	7563	28	05	58	59	84	78	0.871736	20	1	3	-4.5	0.0280	0	1	-0.5862	0.871736	3	0.0883	0.1342	0.1987
7501	7553	52	05	57	58	81	77	0.882606	20	1	3	-4.8	0.0310	0	1	-0.2539	0.882606	5	0.1167	0.1636	0.2244
7501	7592	G3	05	58	59	81	78	0.855282	20	1	3	-4.5	0.0280	0	1	-1.0893	0.855282	2	0.0556	0.0983	0.1674
7576	7564	23	05	60	60	100	80	0.839091	20	2	3	-7.3	0.0245	0	1	-1.5843	0.839091	1	0.0346	0.0716	0.1425
7521	7523	78	05	58	59	81	78	0.895832	20	1	3	-4.5	0.0280	0	1	0.1505	0.895832	6	0.1570	0.2060	0.2654
7511	7592	M2	05	59	59	100	79	0.917163	20	2	3	-6.5	0.0250	0	1	0.8026	0.917163	9	0.2231	0.2904	0.3684
7501	7592	78	05	57	60	100	77	0.849552	20	1	3	-4.8	0.0310	0	1	-1.2645	0.849552	2	0.0472	0.0880	0.1581
7565	7562	22	05	57	57	100	77	0.841952	20	1	3	-4.3	0.0310	0	1	-1.4968	0.841952	1	0.0377	0.0758	0.1466
7522	7530	62	05	57	58	82	77	0.867019	20	2	3	-4.8	0.0310	0	1	-0.7304	0.867019	3	0.0777	0.1220	0.1889
7576	7531	22	05	57	58	83	77	0.884622	20	1	3	-4.8	0.0310	0	1	-0.1923	0.884622	5	0.1225	0.1695	0.2299
7563	7596	22	05	57	59	76	77	0.901605	20	1	3	-4.8	0.0310	0	1	0.3270	0.901605	7	0.1755	0.2269	0.2882
7557	7560	F2	05	61	61	100	81	0.905130	20	1	3	-4.3	0.0265	.	1	0.4363	0.905130	7	0.1868	0.2406	0.3041
7522	180	10	05	60	63	100	80	0.958790	20	1	3	-7.3	0.0245	0	1	2.0753	0.958790	10	0.3351	0.4992	0.6634
7522	9620	N5	05	60	62	100	80	0.921015	20	1	3	-7.3	0.0245	0	1	0.9204	0.921015	9	0.2342	0.3077	0.3925
7522	7560	A2	05	57	58	80	77	0.957610	20	1	3	-4.8	0.0310	0	1	2.0392	0.957610	10	0.3320	0.4929	0.6552
7562	7576	64	05	60	62	100	80	0.905308	20	2	3	-7.3	0.0245	0	1	0.4402	0.905308	7	0.1872	0.2411	0.3047
7501	7557	29	05	57	60	84	77	0.871676	20	1	3	-4.8	0.0310	0	1	-0.5331	0.871676	3	0.0882	0.1340	0.1785

LIEUTENANT COLONEL (AIR) FY81 PREDICTION RATE RESULTS

N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM	VARIANCE
62	0.21324801	0.07636536	0.07637644	0.45643020	0.00969841	13.53137672	0.00583167

LIEUTENANT COLONEL (AIR) FY82 PREDICTION RATE RESULTS

N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM	VARIANCE
49	0.26137732	0.10998338	0.07105955	0.55995518	0.01571191	12.83201337	0.01209634

Figure 4.13 Lieutenant Colonels (Air)

The variables evaluated in this model are shown below in Table XI.

INDEX again emerges as the primary variable for determining attrition for this group and, like the lieutenant colonel (Ground) population group, the coefficient is positive. The other variables analyzed by the model were deleted for the same reasons given during the discussion of the previous data group. Figure 4.14 shows

TABLE XI

VARIABLE SIGNIFICANCE, LIEUTENANT COLONELS (AIR)

Variable	P - Value	Chi - Square
INDEX	0.0001	14.93
UNEMP	0.4113	0.67
PAY	0.6854	0.16
EDSPEC	0.9985	0.00

the final parameter estimates of this data group along with pertinent statistics on the variables deleted from the model.

c. Performance Index Interval Table

Figure 4.15 shows the table for this population group with attrition and retention percentages by interval. Although the total number of officers appearing in this table is much less than the lieutenant colonel ground table, 239 versus 648, the results are similar. Interval 1,

```
                                                   LT COLS (AIR) WITH 20 YEARS
                    STEPWISE LOGISTIC REGRESSION PROCEDURE

                         DEPENDENT VARIABLE: OUT

              239 OBSERVATIONS
              188  OUT      = 0
               51  OUT      = 1
               62 OBSERVATIONS DELETED DUE TO MISSING VALUES

   VARIABLE              MEAN        MINIMUM        MAXIMUM          RANGE

   INDXN             0.891865       0.796998       0.970497       0.173499
   PAY               -5.68703          -7.3           -4.5            2.8
   UNEMP            0.0274142        0.0245          0.031         0.0065
   EDSPEC             1.27615             1              2              1

      -2 LOG LIKELIHOOD FOR MODEL CONTAINING INTERCEPT ONLY=  247.80

STEP  0. THE FOLLOWING VARIABLES ARE ENTERED:

INTERCEP INDXN     PAY      UNEMP     EDSPEC

       MODEL CHI-SQUARE=   16.94 WITH    4 D.F.    (SCORE STAT.) P=0.0020.
       CONVERGENCE OBTAINED IN   5 ITERATIONS.                   R= 0.197.
       MAX ABSOLUTE DERIVATIVE=0.6820D-07.               -2 LOG L=  230.17.
       MODEL CHI-SQUARE=   17.63 WITH    4 D.F.    (-2 LOG L.R.) P=0.0015.

FRACTION OF CONCORDANT PAIRS OF PREDICTED PROBABILITIES AND RESPONSES     :0.681
RANK CORRELATION BETWEEN PREDICTED PROBABILITY AND RESPONSE               :0.382

                          FAST BACKWARD ELIMINATION

   VARIABLE                                     RESIDUAL
   DELETED       CHI-SQUARE          P         CHI-SQUARE       D.F.          P

   EDSPEC           0.00          0.9985         0.00            1         0.9985
   UNEMP            0.67          0.4113         0.67            2         0.7136
   PAY              0.16          0.6854         0.84            3         0.8401

STEP  1.

       MODEL CHI-SQUARE=   16.03 WITH    1 D.F.    (SCORE STAT.) P=0.0001.
       CONVERGENCE OBTAINED IN   5 ITERATIONS.                   R= 0.244.
       MAX ABSOLUTE DERIVATIVE=0.9280D-03.               -2 LOG L=  231.00.
       MODEL CHI-SQUARE=   16.81 WITH    1 D.F.    (-2 LOG L.R.) P=0.0000.

FRACTION OF CONCORDANT PAIRS OF PREDICTED PROBABILITIES AND RESPONSES     :0.686
RANK CORRELATION BETWEEN PREDICTED PROBABILITY AND RESPONSE               :0.391

                  TEST FOR ALL VARIABLES NOT IN THE MODEL
             RESIDUAL CHI-SQUARE=    0.84 WITH    3 D.F.   P=0.8392

RESIDUAL CHI-SQUARE IS NOT SIGNIFICANT AT THE 0.0500 LEVEL.
                                                   LT COLS (AIR) WITH 20 YEARS
                    STEPWISE LOGISTIC REGRESSION PROCEDURE

                         DEPENDENT VARIABLE: OUT

                         FINAL PARAMETER ESTIMATES

  VARIABLE            BETA            STD. ERROR        CHI-SQUARE     P          R

  INTERCEPT   -20.50654636          5.00105997          16.80     0.0000
  INDXN        21.37822017          5.53319721          14.93     0.0001     0.228
```

Figure 4.14 Lieutenant Colonels (Air)

LT COLS (AIR) WITH 20 YEARS

TABLE OF INTERVAL BY OUT

INTERVAL (FREQUENCY / PERCENT / ROW PCT / COL PCT)	OUT	0	1	TOTAL
	0			
INTERVAL 1		21 8.79 91.30 11.17	2 0.84 8.70 3.92	23 9.62
	.847943			
INTERVAL 2		19 7.95 82.61 10.11	4 1.67 17.39 7.84	23 9.62
	.862274			
INTERVAL 3		22 9.21 91.67 11.70	2 0.84 8.33 3.92	24 10.04
	.872510			
INTERVAL 4		20 8.37 90.91 10.64	2 0.84 9.09 3.92	22 9.21
	.881697			
INTERVAL 5		21 8.79 87.50 11.17	3 1.26 12.50 5.88	24 10.04
	.890303			
INTERVAL 6		19 7.95 79.17 10.11	5 2.09 20.83 9.80	24 10.04
	.900135			
INTERVAL 7		23 9.62 92.00 12.23	2 0.84 8.00 3.92	25 10.46
	.910602			
INTERVAL 8		18 7.53 72.00 9.57	7 2.93 28.00 13.73	25 10.46
	.921015			
INTERVAL 9		11 4.60 44.00 5.85	14 5.86 56.00 27.45	25 10.46
	.932495			
INTERVAL 10		14 5.86 58.33 7.45	10 4.18 41.67 19.61	24 10.04
	1.0			
TOTAL		188 78.66	51 21.34	239 100.00

Figure 4.15 Lieutenant Colonels (Air)

officers with a performance index score less than or equal to .847943, shows retention at 91 percent and attrition at 8.7 percent. Intervals 9 and 10, or those officers with the highest performance index, show attrition rates of 56 and 41 percent respectively. Again, contrary to the earlier groups analyzed, this population group shows that officers with the highest performance indexes are attriting at the greatest rate.

8. Colonels (All)

a. Prediction Rates

With this last data group, there were no variables which passed the significance test. Table XII shows the P-value and Chi-Square statistics of the variables deleted from this model. Because these variables did not enter the model, LOGIST assigned the entire population group a uniform probability of leaving the service. This was due in large part to the relatively small size of the population group. The results of this procedure were discarded because they merely equated to a random chance probability of attrition. Accordingly, further analysis of the variables and prediction rates was not conducted for this group.

b. Performance Index Interval Table

The colonels displayed similar results when compared with the lieutenant colonel's ground and air data groups. Attrition is greatest at the highest performance INDEX interval. That interval, interval 10, shows an

TABLE XII

VARIABLE SIGNIFICANCE, COLONELS (ALL)

Variable	P - Value	Chi - Square
PAY	0.0855	2.96
INDEX	0.2001	1.64
UNEMP	0.4870	0.48
EDSPEC	0.5257	0.40

attrition rate of 48 percent for those officers with a score greater than or equal to .933774. The remaining intervals show a substantial decrease in their attrition rates when compared with interval 10. Figure 4.16 shows the results of the interval tables by performance INDEX.

F. SUMMARY

This concludes the data analysis phase of the authors' study of Marine officer attrition. In the following chapter, these results will be consolidated for clarity and the authors will provide comments and recommendations for future studies in this area.

COLONELS (ALL) WITH 26 YEARS

TABLE OF INTERVAL BY OUT

INTERVAL FREQUENCY PERCENT ROW PCT COL PCT	OUT	0	1	TOTAL
	0			
INTERVAL 1		27 7.63 77.14 9.82	8 2.26 22.86 10.13	35 9.89
	.857835			
INTERVAL 2		30 8.47 83.33 10.91	6 1.69 16.67 7.59	36 10.17
	.868889			
INTERVAL 3		27 7.63 77.14 9.82	8 2.26 22.86 10.13	35 9.89
	.877647			
INTERVAL 4		27 7.63 77.14 9.82	8 2.26 22.86 10.13	35 9.89
	.886073			
INTERVAL 5		26 7.34 72.22 9.45	10 2.82 27.78 12.66	36 10.17
	.894349			
INTERVAL 6		28 7.91 80.00 10.18	7 1.98 20.00 8.86	35 9.89
	.903686			
INTERVAL 7		31 8.76 86.11 11.27	5 1.41 13.89 6.33	36 10.17
	.911158			
INTERVAL 8		30 8.47 85.71 10.91	5 1.41 14.29 6.33	35 9.89
	.919439			
INTERVAL 9		31 8.76 86.11 11.27	5 1.41 13.89 6.33	36 10.17
	.933774			
INTERVAL 10		18 5.08 51.43 6.55	17 4.80 48.57 21.52	35 9.89
	1.0			
TOTAL		275 77.68	79 22.32	354 100.00

Figure 4.16 Colonels (All)

V. SUMMARY

A. GENERAL

This chapter will consolidate the results of the output discussed in the previous chapter and offer some comments and recommendations regarding further research in the area of predicting Marine officer attrition. Figure 5.1 depicts, by data group, the prediction sums developed by the SAS LOGIST function along with actual attrition for FY81 and FY82. These figures were transformed into a prediction rate which is also included in Figure 5.1. Additionally, variables which were retained in the model are listed in order of significance while those deleted are presented in a similiar format.

B. PREDICTION RATES

The overall results of this binary choice model, using the SAS LOGIST procedure, were not encouraging. For FY81, three of the five data groups analyzed gave a prediction rate over 90 percent. At first glance, an ability to predict attrition of Marine Officers at a 90 percent or better rate would seem to validate the models and thus provide a useful tool for manpower planning. But, in all three cases, when these same population groups were used to forecast FY82 attrition, the prediction rates dropped well below the 90 percent benchmark achieved for the FY81 data.

POPULATION GROUP	FY81 LOGIST	FY81 ACTUAL	FY81 RATE	FY82 LOGIST	FY82 ACTUAL	FY82 RATE	RETAINED VARIABLES (by significance)	DELETED VARIABLES
1st Lt(GND)	186	183	98%	140	169	83%	INDEX, UNEMP, EDSPEC MOS, PAY	NONE
Capt (GND)	50	51	98%	32	43	74%	INDEX, EDSPEC	MOS, UNEMP, PAY
Capt (AIR)	29	67	43%	7	26	27%	INDEX, PAY, UNEMP, EDSPEC	NONE
Maj (ALL)	*** Deleted from analysis due to insufficient data ***							
Lt Col (GND)	45	34	76%	44	15	34%	INDEX, MOS	UNEMP, PAY, EDSPEC
Lt Col (AIR)	14	13	93%	13	16	81%	INDEX	UNEMP, PAY, EDSPEC
Col (ALL)	*** Deleted from analysis because no variable proved significant ***							

Figure 5.1 Consolidated Results

Additionally, the rates obtained for the remaining data groups were far below the 90 percent level for both FY81 and FY82. Since the results of the tests conducted varied from highly successful to totally unacceptable, the logit procedure would appear to contain sufficient inaccuracy to preclude its use.

This analysis however, has provided further insight into the difficult question of predicting officer attrition. This problem is very difficult to solve using statistical models. Surely, there are a number of variables which cannot be easily quantified yet play a major role in an officer's decision to leave the service. Some of those variables, such as family situation, next duty assignment, personal contacts in the civilian job market and self perception, would have to be individually modeled for this type of analysis to prove statistically relevant to the user. Through personal experiences, the authors have consistently heard the above reasons given as to why an officer has chosen to leave the service. This study seems to confirm the diverse nature of this decision.

Future studies in this area should concentrate on evaluating more population groups with additional variables over a longer period of time to provide a wider range for analysis. Only in this manner can the investigator hope to achieve relevant results for this problem through the use of standard statistical techniques.

C. SIGNIFICANCE OF VARIABLES

As shown in Figure 5.1, INDEX was the most significant variable in each of the five population groups analyzed. These results confirmed the perceptions provided by Major Esmann during his initial research. Officers interviewed for that study felt that promotion potential was a major factor when determining whether to remain on active duty or leave the service. Further, they felt that a high promotion potential would indicate satisfaction with the service while a low promotion potential would cause an officer to resign his commision [Ref. 9]. While the performance INDEX proved significant with all of the population groups, it is of particular note that for the three most senior groups analyzed (lieutenant colonels, Air and Ground, and colonels), the variable had a positive coefficient. The inverse relationship of this positive coefficient means that as an individual's INDEX score increases at these higher ranks, so does the probability that he will leave the Marine Corps. Additionally there is the question of whether the performance INDEX equates to promotion potential. One major shortcoming of this variable in relation to promotion potential was that it failed to reflect an officer's level of experience or his past job assignments. The assignment variable in particular is one of the most significant factors considered by promotion boards and is continually

stressed throughout the Marine Corps as a vital ingredient when planning the "ideal" career. In essence, those Marine officers who have been assigned the "better jobs", may in fact have lower INDEX scores because of the competitive nature of those billets. Nevertheless, these officers will still have a greater promotion potential than their contemporaries who received a higher INDEX score from a less career-enhancing billet. Since the INDEX score has such a narrow range, job assignment may in fact be the vital ingredient for determining promotion potential. A review of the GAO study on quality retention of Marine officers shows that a job assignment fator was numerically transformed and combined with a performance indicator in order to provide a more realistic presentation of promotion potential [Ref. 10]. Future studies in this area should include a job assignment factor to weight the performance INDEX when evaluating the effect of promotion potential on attrition.

While the Esmann study concluded that the unemployment variable(UNEMP) was significant in predicting attrition, the SAS LOGIST procedure provided quite the opposite results. In three of the five data groups analyzed, the unemployment variable did not meet the .05 significance level and was deleted from the model. In those cases where it remained, the population groups were completing their initial contracts and reviewing their employment opportunities in the civilian community. These groups, first lieutenant

(Ground) and captain (Air), were facing similiar decisions and the unemployment rate appears to have been an important factor in their decision to leave the Marine Corps. For the remaining three data groups, the effects of the employment situation were of little or no significance.

During the initial phases of this study it was felt that the educational specialty (EDSPEC) variable would be significant only to the more junior population groups whose recently acquired skills remained a marketable asset. This assumption proved correct, with the first lieutenant through captain data groups retaining EDSPEC and the lieutenant colonel data groups deleting this variable. This would seem to indicate that a college major is more important to a young officer entering the civilian market and less important to the career officer who must sell his managerial background rather than what he studied in college twenty years ago.

The MOS variable was input as a independent variable in only three of the five population groups analyzed. Of those three groups, it was retained by only two as being significant, first lieutenants and lieutenant colonels (Ground). The retention of this variable and its degree of significance in the lieutenant colonels (Ground) model should be noted. It may be that MOS is directly correlated to the job assignment factor previously discussed. It would appear that those officers with combat arms MOS's (0302 and

0802) are in a better position to serve in those assignments considered career enhancing while those officers without these MOS's would not have the opportunity to assume the limited number of command billets. While an individual's past assignments were not introduced into this study, the presence of MOS in this particular model may indicate its relationship with promotion potential and therefore attrition.

The last independent variable to be included in this study was pay comparibility with the civilian sector. For the most part this variable proved insignificant, being retained in only two of the eight data groups analyzed. However, in the case of the captain (Air), PAY was second to INDEX in its level of significance to the model. Although not a principle factor in this population group's decision to leave the Marine Corps, the appearance of PAY in the captain (Air) model still suggests that money is important to this group and may explain the payment of bonuses to retain the younger naval aviator.

To summarize, the viability of the performance INDEX score as a predictor of promotion potential was not determined during this study. However, its value in predicting attrition was confirmed. The same can not be stated of the other independent variables which nevertheless provided some interesting insights into the various elements affecting an officer's decision to leave the Marine Corps.

D. PERFORMANCE INDEX INTERVALS

The performance index interval tables were originally constructed to assist the authors in determining at what INDEX level the greatest attrition was occuring for each population group. Using this information in conjunction with the SAS LOGIST output, the authors would be able to ascertain at which intervals SAS was having the most success in predicting attrition. From the results provided, the indication is that LOGIST predicts quite well at the extreme ends of the tables while the prediction capabilities in the middle intervals are far less accurate.

Through this process the authors discovered that the senior officers, lieutenant colonels - year 20, and colonels - year 26, with the highest performance index scores were attriting at the greatest rate. Table XII shows the retention and attrition percentages at the interval extremes for each of the population groups analyzed. As this table indicates, the percentage of officers staying in or leaving the Marine Corps compared to their performance index interval is completely reversed between the junior and senior officers. If a high performance index equates to being a good officer, then the significance of tnis information would be that the Marine Corps is losing its best officers when they reach the senior ranks. Of course, as discussed previously, the performance index may not

TABLE XIII

CONSOLIDATED PERFORMANCE INDEX INTERVALS

Population Group	Interval	Index Value	Percentage In	Percentage Out
1st Lt (GND)	1	<.796879	47%	52%
	10	>.955157	85%	14%
Capt's (GND)	1	<.843051	83%	16%
	10	>.950598	89%	10%
Capt's (AIR)	1	<.840127	59%	40%
	10	>.944003	91%	8%
Lt Col's (GND)	1	<.838258	83%	16%
	10	>.929695	47%	52%
Lt Col's (AIR)	1	<.847943	91%	8%
	10	>.932495	58%	41%
Col's (ALL)	1	<.857835	77%	22%
	10	>.933774	51%	48%

indicate the total professional ability of Marine Corps senior officers because it fails to consider past assignments when evaluating a Marine's worth to the service. Nevertheless, an officer with twenty years of service has had a variety of assignments with different reporting seniors. If he has been able to maintain an above average performance index score throughout that period, it should be a true indicator of his greater ability and increased promotion potential.

The GAO report used a job experience factor along with a performance indicator which showed that the Marine Corps was

retaining its better senior officers. The opposing results achieved by this study and the GAO report indicates that different methods of evaluating promotion potential will yield varying results. Any future studies in this area should begin by identifying the value of the performance index to determine if it is a true indicator of an officer's performance and, therefore, promotion potential.

E. CONCLUSION

The overall success rate of the logit function to predict Marine officer attrition was less than the authors had expected. A possible explanation may lie with the variables selected to describe the individuals in the population groups. The performance INDEX score is clearly a starting point for determining the promotion potential of an officer, but his past assignments cannot be overlooked as a significant factor in this regard. While the remaining variables provide some interesting insights into the differences between the population groups, their usefulness in predicting attrition was generally minimal. In retrospect, the complexity of this decision, to remain on active duty or leave the Corps, may make it impossible to accurately model using currently available data. Only with the inclusion of a number of additional, difficult to obtain or measure variables, can an analyst hope to achieve accurate results across all population groups and years of service.

LIST OF REFERENCES

1. Center for Naval Analysis 83-3116/22, CNA Memorandum: "Marine Officer Loss Rates for 1984," by Kathleen P. Utgoff, July, 1983.

2. Esmann, William J., Marine Officer Attrition Model, M.S. Thesis, Naval Postgraduate School, Monterey, 1984.

3. "The Pay Gap," Navy Times, 34th Year No. 13, p. 3, 14 January 1985.

4. Esmann, William J., Marine Officer Attrition Model, M.S. Thesis, Naval Postgraduate School, Monterey, 1984.

5. Ott, Lyman, and Hildebrand, David K., Statistical Thinking for Managers, pp. 401-493, Duxbury Press, 1983.

6. Eisenbeis, Robert A., and Avery, Robert B., Discriminant Analysis and Classification Procedures, pp. 1-3, Lexington Books, 1972.

7. Annals of Economic and Social Measurement, "A Comment on Discriminant Analysis "Versus" Logit Analysis," by Danial McFadden, May 1976.

8. SAS Institute, SUGI Supplemental Library User's Guide, SAS Institute Inc., 1983.

9. Esmann, William J., Marine Officer Attrition Model, M.S. Thesis, Naval Postgraduate School, Monterey, 1984.

10. Report to the Secretary of the Navy, "High-Quality Senior Marine Corps Officers: How Many Stay Beyond 20 Years of Service?," General Accounting Office, November 1984.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22304-6145	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93943-5100	2
3.	Professor Dan Boger, Code 54Bk Naval Postgraduate School Monterey, California 93943-5100	1
4.	LtCol David Melchar, Code 0309 Naval Postgraduate School Monterey, California 93943-5100	1
5.	Commandant of the Marine Corps, Code MPP Headquarters, U. S. Marine Corps Washington, D. C. 20380-0001	1
6.	LtCol S. F. Hurst, USMC 9029 Rosewall Court Springfield, Virginia 22152	1
7.	Capt T. E. Manion, USMC 9029 Rosewall Court Springfield, Virginia 22152	1
8.	Maj W. J. Esmann, USMC 4509 Banf Street Annandale, Virginia 22314	1